The Dissertation Committee for Hans Ulrich Grasemann
certifies that this is the approved version of the following dissertation:

A Computational Model of Language Pathology in Schizophrenia

Committee:

Risto Miikkulainen, Supervisor

Ralph Hoffman, Supervisor

Bradley Love

Raymond Mooney

Benjamin Kuipers

Dana Ballard

A Computational Model of Language Pathology in Schizophrenia

by

Hans Ulrich Grasemann, Dipl. Inf.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2010

Acknowledgments

Q: So then couldn't you just fight a snake in lieu of actually writing a thesis?

A: Technically, yes. But in that case the snake would be very big. Very big, indeed.

— LUKE BURNS, The "Snake Fight" Portion of Your Thesis Defense

The snake is defeated, and many people have helped me along the way. I am, of course, just going to assume that they know who they are, and that they are the reason I am emerging from this more or less sane. (What? That's how we do things around here.)

However: My advisors Risto Miikkulainen and Ralph Hoffman deserve much of the credit I am blatantly taking here. I am grateful for their help and guidance. Also, Risto's plan to divide the number of remaining days by the number of unwritten chapters saved this dissertation in the nick of time. Thanks also to the neural networks group at UT Austin (especially Igor) and to my dissertation committee.

The cats and people at the treehouse, the Cd19F, Peggy, Alex, and others all helped in strange and important ways. Without them this dissertation, and the experience of writing it, would have been much less, uh, interesting. Thanks to all of them, and also to Mariella, who saved this dissertation a little more literally than most.

Thanks also to the neo-californians, who provided entertainment and shelter to me and my blank piece of paper. You guys are not very good at bags.

Finally, thanks to my family, who never failed to support me and/or bet me astronomically expensive dinners that I wouldn't be able to finish this by entirely unrealistic deadlines.

The writing of this document involved an exhaustive survey of the coffeeshops in Austin, TX and across northern California, to be published soon in a reputable journal.

ULI GRASEMANN

The University of Texas at Austin

December 2010

A Computational Model of Language Pathology in Schizophrenia

Hans Ulrich Grasemann, Ph.D.
The University of Texas at Austin, 2010

Supervisors: Risto Miikkulainen and Ralph Hoffman

No current laboratory test can reliably identify patients with schizophrenia. Instead, key symptoms are observed via language, including derailment, where patients cannot follow a coherent storyline, and delusions, where false beliefs are repeated as fact. Brain processes underlying these and other symptoms remain unclear, and characterizing them would greatly enhance our understanding of schizophrenia. In this situation, computational models can be valuable tools to formulate testable hypotheses and to complement clinical research. This dissertation aims to capture the link between biology and schizophrenic symptoms using DISCERN, a connectionist model of human story processing. Competing illness mechanisms proposed to underlie schizophrenia are simulated in DISCERN, and are evaluated at the level of narrative language, the same level used to diagnose patients. The result is the first simulation of a speaker with schizophrenia. Of all illness models, hyperlearning, a model of overly intense memory consolidation, produced the best fit to patient data, as well as compelling models of delusions and derailments. If validated experimentally, the hyperlearning hypothesis could advance the current understanding of schizophrenia, and provide a platform for simulating the effects of future treatments.

Contents

List of Figures

Chapter 1

Introduction

Stories are a crucial part of who we are. They enable us to entertain others, to learn from them, and to see the world through their eyes. Stories are deeply informative, and their significance goes beyond communication and social exchange: We make sense of the world and the people around us by fitting our experience into a coherent narrative structure. In schizophrenia, this ongoing narrative breaks down. Disturbances in the perception and expression of reality can be observed through the stories a patient tells. Indeed, narrative language is the primary diagnostic tool, and clinicians use it every day to observe and assess manifestations of schizophrenia. The purpose of cinical interviews, then, is to use narrative language as a window into the schizophrenic mind. The motivation behind this dissertation is the idea that computational models of narrative language can provide mechanistic explanations of what is seen through that window. In other words, this dissertation aims to understand the nature and pathophysiology of schizophrenia as disturbances in computations involved in story processing

1.1 The Biology of Schizophrenia

Over a century ago, Emil Kraepelin, one of the founders of modern psychiatry, asked his assistants Frank Nissl and Alois Alzheimer to look for brain abnormalities in “dementia praecox," the disorder now known as schizophrenia. The cortical and thalamic abnormalities they reported were controversial, but the search for the biological underpinnings of schizophrenia had begun (Bogerts et al. 2009). For much of the 20th century, Kraepelin's view that schizophrenia was likely caused by "a tangible morbid process in the brain" (Kraepelin 1896; c.f. Noll 2006) was overshadowed by the Freudian concept of schizophrenia as a "psychogenic" disease. However, since the 1970s, Kraepelin's theory of genetic and biological defects as key contributors to psychiatric illness has returned to prominence, and today it forms the foundation of biological psychiatry.

Schizophrenia research has come a long way since the days of Kraepelin. It is now clear that schizophrenia is indeed a physical disease, and that structural brain abnormalities, genetic vulnerabilities, and altered brain chemistry are all key components of the disorder. Extensive biomedical research has implicated virtually every brain area and every major neurotransmitter system in schizophrenia (Pearlson and Marsh 1999; Bogerts et al. 2009; Glenthoj et al. 2009). This wealth of experimental findings has led to a large number of hypotheses regarding the causes and pathophysiology of schizophrenia, and recently, plausible accounts of possible interactions between biology, genetics, and symptoms have begun to emerge.

Despite these advances, the precise nature of schizophrenia is still largely a mystery. No laboratory or neuroimaging test can reliably identify persons who suffer from it, and the concept of schizophrenia itself remains a purely diagnostic construct. None of the hypothetical links between biology and symptoms have gained wide-spread consensus, and thus the century-old debate continues (Plum 1972; Ron and Harvey 1990; Kapur 2003).

At the same time, a better understanding of schizophrenia is badly needed, most importantly because the lack of knowledge translates directly into a lack of adequate treatment options. Dopamine-blocking antipsychotic drugs, which have been the mainstay of treatment interventions since the 1950s, are reasonably effective, but address only a subset of the symptoms, do not help all patients, and often cause dangerous side effects (Kane 1997;

Kapur and Mamo 2003). A better understanding of the pathophysiology of schizophrenia would likely lead to more effective drugs, and might suggest entirely new ways to treat or even prevent schizophrenia (Brewer 2005; Pearlson 2000).

Given all the advances in experimental techniques in neuroscience, and given the importance of developing a better understanding of schizophrenia, why has such an understanding has not yet emerged? One major problem is that schizophrenia is highly heterogeneous, i.e. both symptoms and observed biological abnormalities vary widely among patients. This heterogeneity suggests that, rather than a single well-defined illness, schizophrenia is likely to be a family of clinically related disorders that are, at least to some degree, the result of different underlying biological mechanisms. What these pathological mechanisms are, how they co-occur and interact, and how they lead to different patterns of symptoms are all open questions. In short, heterogeneity complicates the question of cause and effect, challenging researchers to infer complex combinations of causes from highly variable patterns of symptoms.

The problem is further complicated by the fact that imaging, neuropathological, and cognitive measures can associate schizophrenia with anatomical, neurochemical, and cognitive changes, but they are in general not able to establish causality. For example, working memory impairment is strongly associated with schizophrenia (Aleman et al. 1999), but it is currently unclear if one is a result of the other, or if they are coincident effects of the same cause. Similar uncertainty exists about many other brain abnormalities in schizophrenia.

In addition to these problems, however, I argue that an even more important obstacle on the way to a better understanding of schizophrenia is the lack of an adequate language in which to express hypotheses about the link between physiology and symptoms in schizophrenia. In other words, new and advanced ways to investigate the schizophrenic brain experimentally should be complemented by new and equally powerful theoretical tools. Contributing to the development of these tools is the principal motivation behind this dissertation.

1.2 Modeling The Brain and Other Complex Systems

The need for new theoretical tools is not unique to schizophrenia research. The traditional formal approach of obtaining analytical solutions to systems of mathematical equations is inadequate for many complex adaptive systems like markets, ecosystems, or the human brain. Until recently, researchers in these areas were limited to verbal theories, which often lack precision and predictive power.

During the last few decades, computational models have emerged as an alternative way to formulate theories, and have gained a prominent role in many scientific disciplines that investigate complex systems, including climate science, economics, meteorology, astrophysics, and many others. Computational models are often based on purely mathematical models, but are distinct from them in important ways. They do not demand a precise description of the system being modeled or its boundaries; they often incorporate other models and span multiple levels of analysis; they are able to approximate the performance of systems that are not well-defined or too complex for analytical solutions; their predictions are explicit rather than implicit and often include unexpected, emergent phenomena.

In cognitive science, computational models are often based on artificial neural networks, where mental processes and behavior are modeled as emergent phenomena in interconnected networks of simple information-processing units. Most neural network-based models are not intended to be physiologically accurate simulations of biological neurons and their interactions in the human brain. Many details of biological neural networks, such as separate neurotransmitter and receptor systems, the complex dynamics of single neurons, and time-dependent information exchange via action potentials, are often either omitted or abstracted into higher level approximations.

Nevertheless, connectionist models tend to exhibit many characteristics of information processing in biological systems, including massively parallel computation, robustness to noise and input errors, and the ability to learn and generalize from limited experience. This property of brain-like information processing in systems that simulate biological mechanisms on an abstract level makes connectionist cognitive models attractive: the simplicity of the underlying framework makes it possible to build models of high-level cognition using mechanisms that are plausible analogs of the real neural substrate.

During the last few decades, neural networks have been used extensively to create models of human cognition in many different domains, including language, learning, vision, and memory (McClelland et al. 1987; Elman 1990; Hinton 1991; Christiansen et al. 1999; Miikkulainen et al. 2005; O'Reilly and Frank 2006). Neural network models such as these have given detailed computational accounts of human behavior and its underlying information processing mechanisms. Moreover, since such models are explicit, running systems, many create unexpexted behavior that will lead to further experimental hypotheses.

1.3 Modeling Schizophrenia

The principal strength of neural network models lies in their intrinsic ability to connect our understanding of biological systems at different levels of abstraction. They bridge the gap between complex mental states and behavior on the one hand and underlying neural information processing on the other. In schizophrenia research, where the central problem is determining the ways in which biological abnormalities lead to altered behavior, this ability is precisely what is needed. A central working hypothesis is that neural networks can be used meaningfully to model not only normal human cognition but also its impairment in psychiatric illness. In other words, neural networks not only function in a brain-like manner, but can also break down in the same way, creating an opportunity to advance our understanding of both the healthy and the disordered brain. This dissertation is motivated by this opportunity. Its goal is to give a computational account of schizophrenia that makes explicit the link between biology and symptoms in terms of altered information processing in a neural network model. Several previous computational studies had similar goals. A number of neural network-based models have been used to simulate research findings related to schizophrenia, including altered working memory (Braver et al. 1999; Cohen and Servan-Schreiber 1992; Monchi et al. 2000), hyperarousal states (Grossberg and Pepe 1970; Grossberg 1999), excessive semantic priming (Spitzer 1997), alterations of functional connectivity between brain regions (Winder et al. 2007), attention (Wang and Fan 2007), impairments of facial affect recognition (Carter and Neufeld 2007), and hallucinations and delusions (Hoffman and McGlashan 1997; Ruppin et al. 1996; Loh et al. 2007).

The approach taken in this work, however, is new and different in several respects. Most importantly, in contrast to previous studies, this research simulates manifestations of schizophrenia using a model of narrative language. This is significant for two reasons. First, conversational language is the primary diagnostic tool used to assess whether a patient has schizophrenia or not. Key symptoms of schizophrenia such as delusions and disorganized speech are observed directly through language, which means that a model of schizophrenic language pathology can be evaluated on a clinically relevant level.

Second, clinical interviews are used to diagnose schizophrenia for good reason: Narrative language and storytelling are among the richest human behaviors, and are considered critical for social intelligence (Bower and Morrow 1990), sense-making (Bruner 1991; Abolafia 2010), and cognition and consciousness (Rubin and Greenberg 2003). Furthermore, narrative language cannot be reduced to the function of a specific brain region or process. Its dysfunction in schizophrenia reveals the disturbance of deeper underlying functions of memory and thought. The purpose of clinical interviews is to use narrative language as a window to the schizophrenic mind, and computational models should be able to do the same.

The modeling work reported here is based on DISCERN, a neural network-based model of human story understanding and recall. The original DISCERN system was introduced by Miikkulainen (1993), and has since been extended to handle complex stories consisting of multiple events (Fidelman et al. 2005; Grasemann et al. 2007). Other extensions in this dissertation include the ability to process emotional context of stories, and the ability to filter distorted output language. DISCERN represents another way in which the approach of this research differs from most previous work. DISCERN is a multi-modular system that combines different network architectures and language-related functions into a single, unified model of human story processing. The complexity of DISCERN translates into a wealth of opportunities to simulate illness mechanisms underlying schizophrenia as network disturbances in the model. This work takes full advantage of these opportunities: A wide range of current research findings were simulated and compared using the model. For instance, the candidate illness models included various disturbances of working memory (Potkin et al. 2009), loss of cortical connectivity (Feinberg 1982/1983), and semantic memory dysfunction (Spitzer 1997), as well as simulations of overarousal and neuromodulatory dysfunction suggested by previous computational models (Grossberg and Pepe 1970; Servan-Schreiber et al. 1990).

Moreover, hyperlearning, a model of aberrant memory consolidation (Grasemann et al. 2009; Hoffman et al. 2010), turned out to be the most important illness model. The hyperlearning/DISCERN model is based on theories of psychosis advanced by Maher (1974) and Kapur (2003), who propose that psychotic symptoms like delusions are due to abnormally intense, or *salient*, experience, possibly driven by dopamine (DA) imbalance. Hyperlearning extends and formalizes this theory by implementing a possible mechanism by which this could occur: Aberrant salience of experience leads to overly intense learning during memory consolidation, which distorts and skews processing of narrative memory, causing psychotic symptoms. This process is simulated in DISCERN using abnormally high network learning rates in language processing modules, and it turns out to have significant predictive power.

The different illness models were applied to DISCERN, and their ability to recreate language abnormalities in schizophrenia was evaluated both quantitatively and qualitatively. The goal of these experiments, and of this dissertation, was to demonstrate that a neural network-based model like DISCERN can be used meaningfully to compare separate illness mechanisms, and to assess their viability as candidate causes of schizophrenia. Illness models can be distinguished by their distinctive language behavior, and can furthermore generate predictions through unexpected, emergent behavior.

Beyond proving the concept, a further goal was to make a contribution to the current understanding of schizophrenia. The need for testable, predictive hypotheses and the attempt to find them are only partly academic. The illness models in DISCERN are detailed computational hypotheses that produce complex behavioral changes at a clinically relevant level. As such, their language behavior can explain and predict clinical data, and has the potential to complement and guide future medical research.

If cognitive science can help advance schizophrenia research, no doubt the reverse is true as well. Our understanding of cognition has greatly improved lately, but global, emergent faculties of the mind like conciousness, intelligence, and personality are still largely a mystery. A mental illness like schizophrenia, where these faculties change and break down in complex ways, offers a unique window into the mind, and an opportunity to investigate how it emerges from its neural substrate.

1.4 Organization

This dissertation is organized as follows. The next chapter reviews the literature on several central topics of this dissertation. The current understanding of story processing in psychology is discussed first, followed by a review of computational models of story processing. The second part of the chapter is dedicated to schizophrenia, including its symptoms and current research findings about its underlying causes. The final section reviews previous computational models of schizophrenia related to this work.

Chapter 3 introduces the computational modeling tools that will be used later to create simulations of schizophrenic language. DISCERN, a connectionist model of human story understanding and recall, is described in detail. Recent extensions of the model are discussed, including the ability to process multi-script stories, a mechanism to attach emotions to story segments, and a filter mechanism that reduces errors at the cost of lower overall language output. Based on the DISCERN model, eight simulations of candidate illness mechanisms that could underlie schizophrenia are then introduced. Each of the illness models is motivated by a specific hypothesis about the pathophysiology of schizophrenia. In each case, the relevant literature is reviewed and the implementation in DISCERN is described.

Chapter 4 describes the steps necessary to develop a set of "healthy" DISCERN systems as a basis for further experiments. An extensive corpus of input stories is designed, semantic word representations are trained using the DISCERN model, and finally, network training methods are optimized and used to produce concrete, running DISCERN systems.

In Chapter 5, the alternative illness models are evaluated experimentally. First, a human subject study of story processing in schizophrenia is described, conducted at Yale as part of a joint project with this dissertation research. The human subject data is used to evaluate quantitatively how well the different illness models are able to match both healthy humans and patients with schizophrenia.

Chapter 6 describes a second computational study that focuses on the more intense symptoms that occur during active stages of schizophrenia. The different illness models are again evaluated and compared, this time focusing on reproducing derailments and delusions, two key symptoms of schziophrenic psychosis.

Chapter 7 discusses and interprets the findings presented in the previous chapters. Strengths and limitations of the modeling approach are evaluated, as well as possible directions for future research.

Chapter 2

Background

This chapter reviews the research literature on the three cornerstones of this dissertation: stories, schizophrenia, and computational modeling. Current understanding of story processing in psychology is discussed first, followed by a review of computational models of story processing. The second part of the chapter is dedicated to schizophrenia. It provides information on its symptoms, emphasizing those expressed through language, and reviews current hypotheses about the link between symptoms and possible causes. The final section of the chapter discusses previous computational models of schizophrenia.

2.1 Stories

Stories are the central theme that binds this dissertation together. The concept of stories, and that of experience related through language, is a key part of our understanding of schizophrenia, and it is also the core of the computational model that forms the basis of this dissertation.

Stories viewed in a narrow sense, i.e. as narrative accounts of real or imaginary events expressed through language, are an important social construct, because they enable us to understand, learn from, and entertain each other. They are the principal tool by which we share experiences. However, they are even more interesting because they reflect and reveal deeper processes through which we interpret, predict, and understand ourselves, our world, and the people in it.

In a more general sense, then, the processes we use to encode ongoing experience can be seen as a kind of narrative behavior, and the resulting memories that represent past (or even fictional) events are stories, even if they are not exchanged with others. The concept that we are living our own personal story is influential in philosophy of mind and consciousness research (Bruner 1991; Rubin and Greenberg 2003). Dennett (1992), for example, in defining the concept of "self", says that

> [...] it does seem that we are all virtuoso novelists, who find ourselves engaged in all sorts of behavior, more or less unified, but sometimes disunified, and we always put the best "faces" on it we can. We try to make all of our material cohere into a single good story. And that story is our autobiography.

Along the same lines, Flanagan (1992) observes:

> Evidence strongly suggests that humans in all cultures come to cast their own identity in some sort of narrative form. We are inveterate storytellers.

Our personal narrative, then, is at least part of what defines us and unifies our identity. But just as importantly, processing stories involves every part of our mind. Memory, attention, problem solving, and social cognition are all intrinsic and intricate parts of the process of encoding, remembering, interpreting, applying, and exchanging stories.

In this sense, the stories we process are only partly linguistic: they are bundles of sensory, interpersonal, emotive, and predictive information as well, and they play many critical roles in our everyday lives. We use stories to give stucture to the events around us. We repeat them to ourselves and to others — not just because they are entertaining or intriguing, but because they are deeply informative. Stories are relentlessly formed, molded and linked in memory because their representations provide the basis for predicting the world, and for understanding the actions and mental states of ourselves and the people around us — a capacity of utmost importance in our complex social world.

Given that story processing is such a rich and complex behavior, it is not surprising that when it is impaired in psychiatric disorders, an equally rich range of symptoms is observed. An important working hypothesis of this dissertation is that disturbances of narrative language allow a deep look at how psychiatric disorders reshape and distort information processing in the brain. A model of altered story processing can then be used meaningfully both to infer deep underlying processes and to test hypotheses about them.

2.2 Script Theory

In order to create a model of altered story processing, it is necessary to start with a model of normal story processing. Such a model must in turn be based on the current understanding of the human psychology involved. *Script theory* (Schank and Abelson 1977) provides such a basis, and forms the fundamental framework underlying the model used in this work.

Script theory models the way humans process stereotypical sequences of events. For example, every time we walk into a restaurant, approximately the same thing happens: we wait to be seated, order a meal and eat it, pay, then leave. The specific restaurant and the people with us may not be the same; the price and quality of the food may change. The basic sequence of events, however, rarely does, and can therefore be learned and reused as a *script*.

Scripts are best understood as templates for certain types of situations, including open slots to be filled in (such as the kind of food), and constraints on what kinds of things can fill the slots (e.g. you cannot order the decor). An instance of a script, then, is a template representation whose slots have been filled to match a specific situation. In order to understand a specific event, all we need to do is find out what kind of script it follows and fill the slots with the appropriate concepts.

Humans use scripts in may different ways to interact efficiently with each other, grasp complex situations, and form expectations about a situation when faced with incomplete information. The restaurant script, for example, would tell us that we are not allowed to pick a table ourselves, unless explicily told to do so. We do not have to ask anybody to bring a menu, and we know without asking that the soup comes first and the dessert last. If events and slot-fillers cannot be observed directly, they can instead be inferred based on prior experience if a script for the current situation is available. Scipts are also used to structure memories efficiently — instead of remembering every detail about an event, the details can be reconstructed later using just a stored script instance and whatever events violated expectations. In fact, when recalling a past experience, humans can often not distinguish observed events from those that were inferred from a script (Graesser et al. 1980).

Script theory was originally conceived as a technique in artificial intelligence, intended as a way to encode procedural knowledge that would enable artificial systems to form expectations and to understand their surroundings in terms of story-like constructs. However, they were soon recognized for their capacity to model aspects of human cognition and language processing, and today scripts are central to the theory of human cognition and memory. The hypothesis that humans use scripts is well supported by experimental evidence. For example, the degree to which events in a story will be remembered can be predicted by whether those events are part of a script (Graesser et al. 1980). Similarly, the amount of time it takes humans to understand a sentence can be predicted by whether it fits into a script (Den Uyl and van Oostendorp 1980). Script theory therefore forms a promising framework for computational models of story processing.

It is important to note, however, that scripts should not be understood simply as a way to process stories efficiently. Rather, script theory provides a general model for the way humans think and learn about the structure of events in their world — narrative language is only one prominent facet of the theory, much like its impairment in schizophrenia is only one facet of deeper defects in the use of knowledge and the flow of information in the brain.

2.3 Computational Models of Story Processing

Script theory was originally created in the context of *contextual dependency* theory (Schank 1972), a model of natural language processing that attempts to represent meaning as a set of discrete tranformations of objects, times, mental states, etc., independently of grammatical structure. Defined in this way, scripts are intrinsically a symbolic concept: Slots are filled by unambiguous concepts; scripts are defined in terms of transitions between physical or mental states, and each script is currently either occurring or it is not. Implementations of script theory were therefore almost without exception symbolic systems. For example, SAM (Cullingford 1978) was the original script-based story comprehension system. Schank (1991) applied the theory to story telling and the idea of an intelligent tutor, and Schank and Cleary (1995) applied the same idea to educational software.

In terms of language and story processing, which was its original goal, script theory was at first not very influential. However, the theory was extended into a model of dynamic memory Schank (1982, 1999). This work was the basis for case-based reasoning (Kolodner 1993), which is still influential in the area of knowledge representation and reasoning.

Recently, scripts have once again attracted interest, this time in the area of artificial intelligence in games (Young et al. 2004; Riedl et al. 2011), intelligent teaching systems (Rowe et al. 2010), and in automatic content generation and the design of interactive media in general (Riedl and Young 2006; Jhala and van Velsen 2009; Riedl 2010). These emerging applications are also motivating new basic research that attempts to understand creative processes and narrative behavior in humans (e.g. Young 2007).

Neural network-based natural language processing has also been an active field in the last three decades. However, in contrast to the abundance of symbolic work on the level of stories and discourse, the focus of connectionist models has been on more traditional problems in computational linguistics, i.e. sentence parsing and representation of semantic knowledge (neural network-based AI in games is an active field, but is not generally based on stories or language; e.g. Stanley et al. 2006).

Many connectionist models of sentence parsing are built on a variation of simple recurrent networks (SRN; Elman 1990), and have modeled many aspects of human sentence understanding successfully (Elman 1990; Miikkulainen 1990; Mayberry and Miikkulainen 2005; Farkas and Crocker 2008; Collobert and Weston 2008). Models of semantic knowledge and lexical access are often based on self-organizing maps (Kohonen 1982), and have been used to model the human bilingual lexicon (Li et al. 2004), large-scale vocabulary acquisition (Sibley et al. 2008), and disorders like dyslexia (Miikkulainen 1997; Plaut 1997) and aphasia (Miikkulainen and Kiran 2009; Grasemann et al. 2011).

In contrast to these sentence- and word-level models of human language processing, the DISCERN model (Miikkulainen and Dyer 1991; Miikkulainen 1993) took a wider view of natural language processing: It was the first integrated subsymbolic system that could model the entire process of human story understanding and recall, from plain-text words to episodic memories, and back to words. Stories in the original DISCERN model each followed a single script, and the task of understanding, remembering, and reproducing them was shared by a collection of cooperating simple recurrent networks and self-organizing maps.

Before we take a closer look at the DISCERN model, the next section provides background information on schizophrenia, including its symptoms and a review of current hypotheses about its underlying pathology. The nature of these hypotheses, combined with the key role of abnormal narrative language in schizophrenia, will illustrate why a subsymbolic model of story processing like DISCERN promises to be well suited to capture central aspects of schizophrenia and its biological causes.

2.4 Schizophrenia

Schizophrenia is a disabling psychiatric disorder characterized by complex alterations in the perception and expression of reality. Patients may suffer from a wide range of symptoms, including hallucinations, bizarre and unusual behavior, delusions, and the inability to communicate effectively via language. The onset of schizophrenia, which usually occurs in adolescence or early adulthood, is often characterized by dramatic psychotic symptoms that tend to wax and wane over time. In later stages of the disorder, these psychotic episodes often give way to other more enduring deficits, including blunted emotions, social withdrawal, and reduced language output.

Schizophrenia is diagnosed by observing a patient's behavior and self-reported experiences. No laboratory test for schizophrenia exists, and in fact underlying brain mechnisms are unknown. Schizophrenia is therefore defined as a collection of symptoms, and while the diagnosis is reliable, its validity can be questioned (Pearlson 2000). Specifically, symptoms and outcome in schizophrenia are highly variable, suggesting that it may not be a single illness but a family of clinically related disorders. This possibility is as real today as it was almost a century ago, when Eugen Bleuler deliberately referred to "the schizophrenias" in the plural when he coined that name (Bleuler 1911).

Despite the fact that its underlying causes are not known, partially effective treatments for schizophrenia exist, mainly based on medication that can help manage psychotic symptoms. Such antipsychotic medication is sometimes considered surprisingly effective given that the neurobiological and neurocognitive basis of schizophrenia is so poorly understood (Kane 1997). However, antipsychotic drugs often have severe and dangerous side-effects, they do not help all patients, and they do not address negative symptoms or cognitive impairment, which can be equally disabling. These shortcomings make better and more targeted treatment an important goal.

This section describes the symptoms of schizophrenia and discusses current hypotheses about the brain mechanisms underlying it, focusing on those most relevant to this dissertation in more detail.

2.4.1 Overview of the Symptoms of Schizophrenia

The manifestations of schizophrenia are complex and span a wide range of altered behavior, perception, and emotion. The following descriptions cover the characteristic symptoms laid out in the current *American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders* (DSM IV-TR, 2000). They are commonly divided into two groups, *positive* and *negative* symptoms.

1. **Positive Symptoms.** These symptoms describe behavior or experiences that are not usually present in healthy individuals, but can be present in schizophrenia. Positive symptoms are often called *psychotic* symptoms, even though the term is less precise, and other, less inclusive definitions exist. In this dissertation, positive and psychotic symptoms are synonyms.

   Delusions are pathological false beliefs that are held despite evidence to the contrary. Not all false or unsupported beliefs are pathological. For example, beliefs that are based on deception or normal religious beliefs do not qualify as delusions (Andreasen 1984). Delusions often share one of several common themes, like the belief that one is a famous person like Napoleon, or that one's thoughts, feelings, or behavior is controlled by outside forces.

   Hallucinations are perceptions that appear real and occur without an outside stimulus. Such perceptions are symptoms only if they occur in a concious and wakeful state. In schizophrenia, auditory hallucinations are very common, often in the form of one or more voices conversing or commenting on the patient's actions.

   Grossly disorganized or catatonic behavior includes inappropriate or bizarre behavior that interferes with regular daily activities. For example, patients may dress inappropriately, or frequently have unprovoked confrontations. Catatonic behavior may include peculiar or absent motor behavior, or lack of reactivity to outside stimuli.

Disorganized speech is fluent spoken language that fails to communicate effectively or follow a coherent discourse plan. It is a manifestation of *positive formal thought disorder*, i.e. it is believed to reflect an underlying impairment of verbal thought. The most prominent signs of disorganized speech are difficulties in maintaining a coherent story line (e.g. *derailment* refers to speech that switches topics without apparent cause), but patients may also show other signs like *blocking* (interruption of speech before it is complete), or (rarely) produce completely incoherent language (*word salad*).

2. **Negative symptoms,** or deficit symptoms, describe the absence of a normal type of behavior or emotional response.

   Blunted affect describes the absence of normal emotional response, both positive and negative. *Anhedonia*, a similar symptom that describes a decrease or absence of feelings of pleasure, is also often mentioned in schizophrenia.

   Alogia, or poverty of speech, is a lack of volume and content of voluntary speech. Patients with alogia tend to give only short answers to direct questions, and sometimes do not say anything without prompting. Alogia is commonly thought to be a sign of *negative formal thought disorder*, i.e. it is assumed to reflect an underlying poverty of thought.

   Avolition is the general absence of drive, motivation and normal goal-directed behavior.

In addition to these symptoms, cognitive impairments in schizophrenia are common but not ubiquitous, including reduced working memory capacity, impaired verbal learning, and disturbed lexical access. These deficits are not used for diagnosis, but they are relatively consistent across the different clinical subtypes, and some researchers believe that cognitive impairment (especially working memory impairment) should be regarded as a central pathology in schizophrenia (Lewis and Gonzales-Burgos 2006). The true significance of cognitive impairment is currently unclear, but it is by no means specific to schizophrenia (Crockett et al. 1988), and could also be the result of the stress involved in such a severe psychiatric illness (Arnsten and Goldman-Rakic 1998).

Based on commonly co-occurring patterns of symptoms, five clinical subtypes of schizophrenia have been identified, including the *paranoid type*, where symptoms include delusions and hallucinations but not prominent thought disorder, and the *disorganized type*, where symptoms are dominated by prominent thought disorder, disorganized behavior, and emotional blunting. The other three types are *catatonic*, *undifferentiated*, and *residual* type schizophrenia.

2.4.2 Diagnosis

Patients with schizophrenia rarely have all of the above symptoms, and patterns of symptoms vary widely from patient to patient. In addition, most symptoms are shared with other illnesses. Psychosis, for example, occurs in a variety of other disorders, including bipolar disorder, drug abuse, clinical depression, and other psychiatric or medical conditions. Delusions also occur in delusional disorder. Negative symptoms overlap with the symptoms of depression; social withdrawal can be a sign of anxiety disorder. These and other similarities make the differential diagnosis of schizophrenia difficult.

Several standardized diagnostic criteria exist for schizophrenia that are intended to make consistent and reliabe diagnosis possible. For example, the criteria that are currently used in the United States, as well as for most research studies, are those described in the DSM IV-TR. In order to be diagnosed with schizophrenia according to these criteria, a patient must only display two of the seven symptoms, at least one of them positive. If hallucinated voices are present that comment on the patient's life, or if delusions are judged to be bizarre, only that symptom is required. However, in all cases, the patient must suffer from social or occupational dysfunction, and signs of the disorder must persist for at least six months.

The diagnosis of schizophrenia relies almost entirely on language: A clinical interview is conducted with the patient, and diagnosis is made based on the observed behavior. Information comes from two primary sources: The symptoms reported by the patient, and the signs observed through his or her language. Hallucinations, for example, would be diagnosed because the patient says something like "I hear voices," not because the interviewer is able to hear them. In contrast, disorganized speech is observed directly in the discourse structure and in abnormal speech patterns. Similarly, delusions can be observed directly, for example when a patient offers bizarre opinions or states patently false things as fact. Alogia can also be directly observed when patients talk very little and volunteers no information in a conversation.

The focus of this research is on symptoms that can be observed directly in the language of schizophrenic patients, most importantly disorganized language and delusions. The following sections discuss these symptoms in more detail.

2.4.3 Disorganized Speech

Language disorganization can take a variety of different forms. Patients with schizophrenia often have problems maintaining a consistent storyline. For example, *circumstantiality* describes extreme long-windedness or speech that is delayed by unnecessary or irrelevant detail. *Distractable speech* means the patient changes subjects easily, but in response to a real stimulus (i.e. not without apparent cause). The following example was taken from the *Scale for the Assessment of Positive Symptoms* (SAPS; Andreasen 1984):

> Then I left San Francisco and moved to ... where did you get that tie? It looks like it's left over from the 50's [...]

Derailment is a speech pattern where a patient switches from one topic to another that may be only vaguely related or completely unrelated to the current one, leaving the listener in a bewildered state. Even if no single derailment is particularly severe, steady slippage can lead to answers that have nothing in common with the original question. Derailed language in schizophrenia often seems disjointed and fragmented, giving the impression of an arbitrary juxtaposition of "discourse shards." The following example is part of an interview with a thought-disordered patient (Andreasen 1984):

> *Interviewer:* Did you enjoy college?
>
> *Subject:* Um-hum. Oh hey well, I really enjoyed some communities I tried it, and the, and the next day when I'd be going out, you know, um, I took control like uh, I put, um, bleach on my hair in, in California. My roommate was from Chicago, and she was going to the junior college. And we lived in the Y.M.C.A., so she wanted to put it, um, peroxide on my hair, and she did, and I got up and looked at the mirror and tears came to my eyes. Now do you understand it, I was fully aware of what was going on but why couldn't I, I... why, why the tears? I can't understand that, can you?

In extreme cases, disorganized language can be entirely incomprehensibe at times. The main difference between such "word salad" and severe derailments is that the breakdown occurs *within* the sentence structure, which is not the case for derailments. This type of behavior is rare in schizophrenia – in general, the sentence structure of thought-disordered schizophrenic patients is intact.

Apart from these difficulties in following a coherent discourse plan, thought-disordered patients sometimes repeat words or phrases (*perseveration*), make word choices based on rhymes and puns rather than meaning (*clanging*), or stop speaking before a thought is completed (*blocking*). Primarily, however, disorganized language in schizophrenia is characterized by a break-down at the level of overall discourse structure, not at that of words and sentences.

2.4.4 Delusions

Delusions can take many different forms, but in almost all cases, they follow one of a surprisingly small number of themes, for example:

- **Grandiose delusions** include the belief that one is a famous person, for example a rock star or Jesus Christ, or believing that one has special powers or abilities.

- **Persecutory delusions** are beliefs that one is being conspired against or persecuted. Examples include secret agents being after the patient, coworkers or neighbors harassing the patient, the phone being bugged and the mail opened, etc. Persecutory delusions can be extremely complex and self-consistent.

- **Delusions of reference** are delusions where the patient thinks that insignificant events refer to him or have special significance. Items on the news, for example, may be seen as messages intended specifically for the patient.

- **Delusions of thought insertion, thought control, thought broadcasting.** These are delusions where the patient thinks his thought are controlled or inserted by an outside force, or that his thoughts are being broadcast so that he or others can hear them.

- **Delusions of control** are the experience of one's actions and feelings being controlled by an outside force. For example, the patient may feel that aliens are controlling his brain with radio waves.

Apart from the content, delusions are generally classified as *bizarre* (e.g. aliens removing the patient's brain) or non-bizarre (being followed by the CIA). The severity of a delusion is also judged by its complexity and self-consistency, and by the patient's ability to question his beliefs.

Delusions in schizophrenia are often bizarre and frequently have paranoid content. Patients with persecutory delusions tend to confuse the actors and agents in their personal stories with those of the shared stories of their culture. Often, patients inserts themselves into stories that are imaginary or unrelated to the patient. Such “agency shifts” are thought to be the cause of the spurious plots and imaginary conspiracies that characterize persecutory delusions.

The common themes of delusional beliefs change to some extent over time, and adapt to cultural context. For example, unsurprisingly, delusions today include witches much less frequently than a century ago. On the other hand, delusions of being controlled seem to have occured in a stable fraction of schizophrenia patients between 1886 and 1946 (Kranz 1967). Such shared patterns in delusional beliefs suggest that they have a common underlying cause, and that the same essential stories are adapted to the individual and his cultural context in each case. The DSM-IV defines a delusion as

> A false personal belief based on incorrect inference about external reality and firmly sustained in spite of what almost everyone else believes and in spite of what constitutes incontrovertible and obvious proof or evidence to the contrary [...]

Maher (2002) pointed out several ways in which this definition is problematic. For example, many delusions, like being followed by secret agents, are actually possible, and many others can only be refuted with extreme difficulty. More importantly, the definition contains the assumption that delusions are caused by incorrect inference. In fact, experimental data suggest that formal, syllogistic reasoning in delusional patients is not impaired compared to healthy individuals (Kemp et al. 1997). Furthermore, healthy persons frequently acquire beliefs by incorrect inference, and indeed often hold irrational beliefs like having seen UFOs (Gallup and Newport 1991). Note that in this case as well, proving incontrovertibly that no UFO has ever visited the Earth would be impossible.

If the abnormality underlying delusions is not incorrect inference, then, what other cause could there be? Maher (1974, 2002) proposed that some delusional ideas are based on anomalous experience. Simply stated, a failure to predict events leads to a feeling of undirected significance, which causes the need for an explanation. Finding such an explanation, even a delusional one, brings relief, and evidence to the contrary is ignored to protect the explanation. Maher also pointed out that not all delusions need to be formed in the same way, just like non-delusional beliefs may be acquired in many different ways.

2.5 Possible Causes of Schizophrenia

After many decades of research into the causes of schizophrenia, not much is known with any certainty. Advances in neuroimaging, pharmacological, and post-mortem techniques have at least demonstrated one fact: Schizophrenia is a physical illness, involving genetic, developmental, anatomical, and neurochemical abnormalities (Weinberger 1995). The exact nature of these abnormalities, which ones are central, and how they cause symptoms, are still open questions, and the debate continues (Plum 1972; Ron and Harvey 1990; Kapur 2003).

This section discusses some current attempts to answer these questions. In order to manage the vast research literature on the subject, I will focus on hypotheses about schizophrenia that establish a link between biological processes on the one hand and abnormal experience and behavior on the other. Work that deals with only one of the two is only included if it provides converging evidence.

When discussing specific hypotheses about schizophrenia, it is important to keep in mind that schizophrenia is heterogeneus, which makes the search for underlying illness mechanisms much more complicated, and opens up a number of additional questions: How many different illnesses are there? Do they correspond to known clinical subtypes? Which symptoms arise from which underlying mechanisms, and why do they tend to occur together? The problem has become one of inferring many different causes from many effects, and, as Maher and Deldin (2002) put it,

> Arguing backwards from effect to cause is notoriously unreliable, but it is a constant problem for the research pathologist who takes heterogeneity seriously.

Despite these difficulties, a number of plausible partial explanations have begun to emerge. The remainder of this section reviews some of them, emphasizing the ones that are relevant to this dissertation.

2.5.1 Abnormal Brain Connectivity

In schizophrenia, both local cortical connections (microcircuitry) and connectivity between different regions of the cortex (macrocircuitry) were found to be altered (Bogerts et al. 2009). These alterations are especially interesting because schizophrenia has long been considered a developmental disorder (Weinberger 1987; Feinberg 1982/1983), and humans ordinarily lose about 40% of their synapses through a developmental process called *cortical pruning* (Huttenlocher 1979; Chechik et al. 1998; Abitz et al. 2007) — a process that ends during the same developmental period in which schizophrenia usually begins, namely late adolescence to early adulthood. This suspicious timing is the original motivation behind the *pruning hypothesis* of schizophrenia, which suggests that overeager or prolonged developmental pruning contributes to the emergence of schizophrenia (Feinberg 1982/1983; Saugstad 1994).

The pruning hypothesis has received further support from several studies indicating that neuropil (the tangle of axons and dendrites surrounding cortical neurons) is reduced in schizophrenia. Imaging studies have shown less gray matter in patients with schizophrenia than in normals (Buchanan et al. 1998), especially in the prefrontal cortex (PFC), a brain area implicated in many symptoms of schizophrenia, and also one where developmental pruning is especially pronounced (Huttenlocher 1997). Furthermore, neurons are more dense than normal in several prefrontal areas in schizophrenia (Selemon et al. 1995, 1998), while the actual number of cortical neurons is not changed (Pakkenberg 1993). Taken together, these findings suggest that neuropil is reduced and that the cortex is less well connected in schizophrenia.

Recently, advances in functional brain imaging and EEG analysis have revealed another kind of abnormal connectivity in schizophrenia. Several researchers reported patterns of changed functional connectivity between brain regions in schizophrenia. For example, Kim et al. (2003) observed functional disconnection between prefrontal and parietal cortices during working memory processing in schizophrenia in a PET study. More recently, Karlsgodt (2008) reported anatomical changes in frontal-parietal white matter connections and linked such changes to performance on working memory tasks. Another imaging study (Meyer-Lindenberg et al. 2001) found similar disruptions of interactions between brain regions, and Sakkalis et al. (2006) observed disconnections using EEG data. These findings support the view that schizophrenia may be a disorder of brain connectivity and involve white-matter pathology.

Interestingly, both local PFC connectivity and frontal-parietal networks are thought to play key roles in working memory function (Smith and Jonides 1997; D'Esposito et al. 2000; Postle et al. 2003). In schizophrenia, working memory, along with other executive functions associated with the frontal cortex, is definitely impaired (Nuechterlein and Dawson 1984; Bilder et al. 2000; Gur et al. 2007), and PFC function in general is often considered central to the pathophysiology of schizophrenia (Bogerts et al. 2009).

2.5.2 Dopamine

In its original form, the *dopamine hypothesis* of schizophrenia suggests that some symptoms of schizophrenia are caused by abnormal dopaminergic transmission in the brain, probably involving overactivity of midbrain dopamine (DA) neurons that project to limbic and cortical regions. It originated with the observation that antipsychotic drugs are effective because they block the D2 dopamine receptor (Kapur and Seeman 2001), while DA agonists on the other hand can provoke psychotic symptoms even in healthy individuals (Egan and Weinberger 1997; Curran et al. 2004), and are much more likely to do so in psychotic patients (Lieberman et al. 1987). Imaging studies have provided further evidence for excessive DA release in schizophrenia, especially when patients are having active symptoms (Guo et al. 2003; Laruelle 2000).

On balance, the evidence that abnormal DA activity is involved in schizophrenia in some form is strong. However, there is some uncertainty whether or not it is a direct cause of symptoms. First, active schizophrenia is a very stressful condition (Dazzi et al. 2004), and stress is known to elevate midbrain DA release. Second, a number of studies have suggested reduced cortical DA release in patients with schizophrenia (Lidow et al. 1997; Goldman-Rakic et al. 2004). Third, blocking DA completely through drugs does not universally or immediately remove active symptoms.

Until recently, detailed accounts of how DA imbalance could lead to manifestations of schizophrenia were rare. The complex functions of DA in the brain, including in reward and reinforcement, synaptic plasticity, and working memory were not sufficiently understood. However, advances in the understanding of DA function have led to several plausible theories, and today the DA hypothesis subsumes a number of different accounts of links between DA and schizophrenia.

One widely endorsed emerging theory is due to Kapur (Kapur 2003; Kapur et al. 2005). Based on the view that DA activity mediates *motivational salience* (Berridge and Robinson 1998), and extending Maher's model of delusion formation (Maher 1974, 2002), Kapur proposes that in schizophrenia, increased midbrain DA release leads to abnormally enhanced motivational salience. The salience in turn is thought to cause psychotic symptoms through normal mechanisms of memory, learning and inference, but based on abnormal and overly intense experience. Delusions, for example, are explained as secondary reactions to an altered experience of the world — i.e. trying to make sense of the aberrant significance assigned to insignificant events or facts. Hallucinations are simply seen as normal percepts and memories that are pathologically enhanced by abnormal salience.

A different take on the DA hypothesis was prompted by computational models of the role of DA in working memory (WM; Servan-Schreiber et al. 1990; Trantham-Davidson et al. 2004). In particular, Braver et al. (1999) offered a possible explanation of how abnormal DA transmission in the midbrain and frontal cortex might lead to cognitive deficits observed in schizophrenia. According to this theory, *tonic* DA activity (possibly mediated by the D1 receptor) serves to stabilize activation patterns stored in working memory, while *phasic* (D2) dopamine activity temporarily destabilizes those patterns to facilitate updates. Thus, DA imbalance could make working memory either unstable or too stable, possibly explaining cognitive impairment and negative symptoms in schizophrenia. Additionally, overactivity of midbrain D2 dopamine could lead to faulty information gating and spurious working memory updates without meaningful input, providing a possible model for psychotic symptoms. Recent reports of excessive cortical noise, reduced signal-to-noise ratio, reduced control, and reduced efficiency in frontal cortical systems linked to WM in persons with schizophrenia Potkin et al. (2009); Tan et al. (2007); Winterer and Weinberger (2004) support the view that WM dysfunction plays a role in schizophrenia.

The two versions of the dopamine hypothesis described above — abnormal working memory function and aberrant motivational salience — largely concern different complexes of symptoms and are formulated at different conceptual levels. Nevertheless, a classic theory first proposed by Weinberger (1987) suggests a possibility to unify the two. Weinberger postulates a "paradoxical state" of DA function in schizophrenia, where low cortical DA levels and high midbrain DA levels coexist. Combining the three hypotheses, low prefrontal DA activity could account for cognitive deficits and negative symptoms, while high midbrain DA would produce heightened phasic D2 activity, a possible neural substrate of aberrant salience. According to Weinberger, such a paradoxical state could be caused by a primary impairment of prefrontal DA function wich could cause a chain reaction involving cortical, midbrain, and limbic systems that induces excessive D2 midbrain activity. However, the opposite direction of causality is also possible: A recent animal study by Kellendonk et al. (2006) suggests that even temporarily increased midbrain DA transmission can cause persistent abnormalities of prefrontal DA function and WM impairment.

In more recent version of the DA hypothesis, Assadi et al. (2009) proposed that dopaminergic dysregulation of a network involving dorsal anterior cingulate cortex (dACC), striatum, and thalamus could contribute to the pathophysiology of schizophrenia. The dACC and its related subcortical structures are involved in decision making (Kennerley et al. 2006; Hampton and O'Doherty 2007), and the midbrain DA system regulates this network (Salamone et al. 2007). According to this hypothesis, failure of the decision-making process leads to negative symptoms and disorganization in schizophrenia. The faulty decision network fnord does not cause delusions and hallucinations, but delusions and halluciananations cause social dysfunction because of impaired decision-making.

In summary, there is strong evidence that DA is involved in schizophrenia, but there is no consensus yet as to how DA dysregulation could cause symptoms. Given the wide range of DA functions in the brain, DA imbalance could contribute to the manifestations of schizophrenia in many different ways. Emerging hypotheses are often not mutually exclusive, and include working memory impairment, aberrant motivational salience, and dysregulated decision-making networks.

2.5.3 Impaired Semantic Memory

Converging evidence from imaging, psycholinguistic, computational, and lesion studies suggests that map-like cortical structures encode semantic knowledge in the human brain (Farah and Wallace 1992; Caramazza et al. 1994; Spitzer et al. 1995). Semantic concepts in such topographic maps are localized, and the map structure is organized according to similarities in the use and meaning of concepts.

Given this development in the understanding of associative semantic memory, Spitzer (1997) argued that early characterizations of schizophrenic pathology in terms of association psychology become significant again. Specifically, the classic observations that free word associations produced by schizophrenic patients tend to "proceed along new lines" and that “indirect associations receive unusual significance” (Bleuler, 1911, c.f. Spitzer, 1997) suggest that semantic maps, or access to semantic maps, are disturbed in schizophrenia.

Disturbed lexical access in schizophrenia has been investigated recently using lexical decision tasks, where subjects decide whether a given string of characters is a word or not. A robust phenomenon demonstrated in both healthy and schizophrenic subjects is *semantic priming*: a target word is recognized faster if it is preceded by a closely related word called a *prime* (e.g. black → white). This effect has been attributed to activation in the semantic map that spreads from the prime to related words and pre-activates the target word (Neely 1977, 1991). Several studies demonstrated that this effect is more pronounced in thought-disordered (TD) schizophrenic patients (Manschreck et al. 1988; Kwapil et al. 1990; Spitzer et al. 1993a), although those findings have not been universally replicated (Chapin et al. 1989; Ober et al. 1995; but see Moritz et al. 2003). Even greater increases in priming for TD schizophrenic patients were observed when the relation between prime and target word was indirect, as in black → (white) → snow, or lion → (tiger) → stripes (Spitzer et al. 1993a).

The increased priming effect, combined with the map-based model of semantic memory, has led to the *hyper-priming* hypothesis, which states that excessive spreading activation in semantic maps is a major contributor to the symptoms of thought disorder in schizophrenia (Maher et al. 1987; Spitzer et al. 1993b; Aloia et al. 1998; Moritz et al. 2003). The indirect priming effect suggests furthermore that the activation spreads not only faster, but also farther in TD patients with schizophrenia (Spitzer et al. 1993a; Spitzer 1997).

Spitzer also argued that a combined dysfunction of semantic and working memory can account for other clinically relevant symptoms as well, including the pervasive lack of sensitivity to context found in schizophrenia patients. He also pointed out that the observed differences in lexical access could be due to deeper structural or functional cortical deficits. Indeed, studies of verbal fluency and object comparisons suggest that semantic memory itself, not just access to it, may be disorganized in TD schizoprenic patients (Goldberg et al. 1998; Tallent et al. 2001). This hypothesis is also supported by the fact that semantic priming abnormalities are independent of medication, and seem to persist over the entire course of schizophrenia (Moritz et al. 2003). To complicate things further, imaging studies have shown cortical overactivation during semantic association tasks (Kuperberg et al. 2007; Assaf et al. 2006), suggesting that overpriming effects may be due to general semantic overactivation.

In summary, abnormal lexical access in TD schizophrenic patients has prompted the hypothesis that excessive spreading of activation in map-like semantic networks is a major contributor to thought disorder in schizophrenia. Disturbances of semantic memory in TD schizophrenic patients are well established, but their exact nature, and whether or not they cause thought disorder, is currently unclear.

2.5.4 Other Possible Causes

In addition to the hypotheses about the causes of schizophrenia discussed so far, numerous others have been proposed. Structural abnormalities in many different parts of the brain have been linked to schizophrenia, including in the hippocampus, basal ganglia, PFC, cingulate cortex, and others (see Bogerts et al. 2009 for a review).

Apart from dopamine transmission, many other biochemical alterations have been reported, implicating almost every major neurotransmitter system, including serotonin, GABA, glutamate, norepinephrine, and others (see Glenthoj et al. 2009 for a review).

Historically the *immune hypothesis* of schizophrenia has received much attention, which states that neuroinflammatory (possibly autoimmune) processes are significant in schizophrenia (Bogerts et al. 2009). Another influential hypothesis is due to Crow and colleagues, who state that abnormal brain development (possibly including connection pruning) could lead to a failure to achieve left or right dominance, which could ultimately cause schizophrenia. (Crow 1990). Postmortem and imaging studies have confirmed that schizophrenic brains are sometimes less lateralized, and schizophrenic patients are more likely to be left-handed or mixed-handed than the general population (see Maher 2002 for a review).

More recently, a mechanism that has been implicated in schizophrenia is apoptosis, i.e. programmed cell death that is regulated by a complex mechanism of cell signals. Recent postmortem studies suggest that apoptotic mechanisms in several cortical regions are dysregulated (Glantz et al. 2006). Furthermore, certain pro-apoptotic triggers can lead to a "sub-lethal" form of apoptosis that eliminates synapses without neuronal loss, a mechanism that could account for the reduced neuropil (LF et al. 2005; Glantz and Lewis 2000).

In summary, there is evidencethat many different neurochemical, cognitive, and anatomical abnormalities are involved schizophrenia. Competing hypotheses, which are not always mutually exclusive, often involve dopamine imbalance, developmental abnormalities, or disturbed semantic memory. The heterogeneity of schizophrenia makes the search for underlying causes more difficult, and it is unlikely that any one theory will be able to explain the complex cognitive, emotional, and behavioral changes in schizophrenia. Given the daunting nature of this task, new tools to formalize and test hypotheses and to complement and guide clinical research are needed.

2.6 Computational Models of Schizophrenia

During the last two decades, neural network-based models have increasingly been used in research that attempts to capture and express central aspects of psychiatric and neurological disorders in a computational process. Models of many known disorders have been proposed, including conditions as varied as Alzheimer's disease (Finkel 2000; Adeli et al. 2005), epilepsy (Wendling 2008), depression (Huys 2007), dyslexia (Harm and Seidenberg 1999; Miikkulainen 1997) and schizophrenia (Grossberg and Pepe 1970; Hoffman et al. 1986; Cohen et al. 1996; Reggia et al. 1999; Spitzer 1997).

These models are useful because they can potentially link underlying causes of brain disorders to their behavioral manifestations, and because they can express hypotheses about such links in a formal yet flexible and predictive way. This ability makes them attractive theoretical tools to complement experimental research. Schizophrenia has been an important focus for computational models of psychiatric disorders, both because of its complexity and because experimental research has so far failed to identify its causes. Schizophrenia thus presents an opportunity to demonstrate the value of computational models as well as a challenge to create testable computational hypotheses of such a complex and heterogeneous disorder.

As early as 1970, based on the theory that schizophrenic patients are in a constant state of overarousal, Grossberg and Pepe (1970) described a neural network model where pathological reductions of spiking thresholds caused an increased span of learned associations, suggesting a model for attention deficits and loose associations in schizophrenia. Later, Grossberg (1999) developed a model of hallucinations based on Adaptive Resonance Theory (ART) where hyperactive modulatory signals intensify top-down predictions such that perceptions occur without external input. Again based on ART, Grossberg (2000) advanced a model that attempts to explain negative symptoms in schizophrenia as imbalanced opponent processes between emotional centers and the cortex.

PARRY, another early simulation of schizophrenia, was introduced by Colby (1973). PARRY was a symbolic system rather than a neural network-based model, but it is especially relevant to this dissertation because it attempted to simulate the language of a paranoid schizophrenic. The system included a behavioral model based on conceptualizations and beliefs, i.e. judgments about specific conceptualizations. Paranoid schizophrenia was modeled as a dysfunction of the belief system, leading to abnormal judgments to accept or reject propositions.

In another early study of psychopathology, Hoffman (1987) investigated perturbations in attractor networks (Hopfield 1982) to model dysfunctional associative memory in schizophrenia. In response to information overload in such networks, instabilities and "parasitic" stable states emerged, suggesting disorganization and delusions in schizophrenia.

Several other models have also focused on the positive symptoms of schizophrenia. Ruppin et al. (1996) studied a computational model of Stevens's (1992) theory of the pathogenesis of schizophrenia. This theory suggests that in schizophrenia, projections from the medial temporal cortex to the frontal cortex are lost, and symptoms then emerge because synapses regrow at the projection sites to compensate for the lost connectivity. Ruppin and colleagues simulated these hypothesized synaptic changes in a model of the frontal cortex based on an attractor network. They observed spontaneous, stimulus-independent retrieval of stored memories that focused on just a few of the stored patterns, suggesting hallucinations and delusions that occur in schizophrenia without apparent external cause.

In a more recent study, Loh et al. (2007) used similar but more biologically detailed networks to model symptoms of schizophrenia through changes in network dynamics. Reducing the depth of basins of attraction in spiking attractor networks reduced the stability of memory states, suggesting a cause for cognitive and working memory deficits. Lower firing rates were observed that could account for negative symptoms, and spontaneous jumps into attractor states could explain positive symptoms.

In a study closely related to this dissertation, Hoffman and McGlashan (1997) used simple recurrent networks (of Elman 1990) to model aspects of human speech perception. In order to understand the mechanisms underlying hallucinated speech in schizophrenia, the impact of several simulated pathologies on speech perception was investigated. When recurrent connections were pruned excessively, the networks generated spontaneous speech percepts, thus emulating hallucinated speech. A further study (Hoffman and McGlashan 2006) compared the performance of "hallucinating" networks to that of actual hallucinating patients. Overpruned networks with additional simulated downregulation of dopamine activity matched human data best, suggesting that schizophrenia may arise from curtailed connectivity and involve secondary downregulation of dopaminergic activity.

In contrast to these simulations of psychotic symptoms, Cohen and colleagues (Servan-Schreiber et al. 1990; Cohen and Servan-Schreiber 1992; Braver et al. 1999) focused on modeling behavioral deficits and cognitive impairment in schizophrenia. They argued that these manifestations of schizophrenia can be explained by a failure of cognitive control and processing of context due to faulty information gating in working memory. The gating failure in turn was thought to be caused by abnormal neuromodulatory influence, especially through DA. The theory was implemented in a series of connectionist models of cognitive control tasks like the Stroop test. The influences of simulated abnormal DA activity were measured, and shown to account for cognitive deficits in schizophrenia.

Based on the hypothesis that schizophrenic thought disorder is caused by dysfunctional map-like semantic networks (Section 2.5.3), Spitzer (1997) proposed a neural network-based model of disturbed lexical access in schizophrenia based on self-organizing maps (SOMs, Kohonen 1982). In this model, access to semantic memory is impaired through excessive spreading activation, and it is argued that activations in SOMs that spread faster and farther than normal can account for increased indirect priming effects, decreased accuracy of lexical access, and other correlates of thought disorder in schizophrenia.

In a related study,Silberman et al. (2007) used a SOM-based model of semantic memory to simulate how semantic and episodic factors interact to form word associations. In this model, activations that spread faster and farther than normal were simulated in order to gain insight into the causes of schizophrenic thought disorder. The results of this study suggested that impaired spreading activation may indeed be able to account for impaired associative thinking in thought-disordered schizophrenia patients.

Carter and Neufeld (2007) investigated impairments of facial affect recognition in schizophrenia using a partially recursive network trained with backpropagation through time. Several competing impairment models, including connection pruning and altered network gain, were evaluated using human subject data. The hypothesis that additional network processing load interferes with the judgment of facial affect provided the best match for the impairment seen in schizophrenia.

The computational studies discussed above span three decades, and simulate a wide range of symptoms and underlying illness mechanisms using a variety of different approaches and network architectures. They all represent progress towards a true "computational patient." One problem shared by most of them, however, is that manifestations of schizophrenia are represented by network behavior that is either very abstract, as is the case of attractor networks, or does not correspond directly to core symptoms, as in the simulations of cognitive defects or facial affect recognition. In both cases, the link between network behavior and schizophrenia is subject to interpretation. The model presented in this dissertation attempts to avoid this problem by simulating symptoms on the level of narrative language, the same level used to define and diagnose schizophrenia.

2.7 Conclusion

The purpose of this chapter was to provide the background on three overarching themes of this dissertation: stories, schizophrenia, and computational modeling. One important goal of this research is to combine these three themes, and to demonstrate that a computational model of human narrative language can also simulate how language breaks down in schizophrenia. An accurate simulation of impaired story processing could be expected to reveal deeper distortions of information processing in the brain, and could then be used meaningfully to infer underlying illness mechanisms. The next chapter introduces a set of computational tools that will be used for this purpose.

Chapter 3

Computational Modeling Approach

The goal of this dissertation is similar to the goals of the computational studies discussed in the previous chapter: To develop a computational understanding of schizophrenia, specifically of the ways in which abnormal brain processes can cause symptoms to emerge. However, the approach taken in this work is new and different in several respects.

Most importantly, this research attempts to simulate manifestations of schizophrenia at the level of narrative language — the same level at which real patients are diagnosed. The symptoms of schizophrenia span a wide range of altered behavior and perception, including bizarre behavior and social dysfunction. However, since schizophrenia is diagnosed mainly through clinical interviews, the most relevant human behavior is conversational language. For example, disorganized speech and delusions, two hallmark symptoms of schizophrenia, are observed directly in the patient's conversational language (rather than the patient reporting symptoms via language). Moreover, as I argued in the previous chapter, narrative language is one of the richest human behaviors. It cannot be reduced to the function of a specific part of the brain, and its dysfunction in schizophrenia has the potential to reveal the disturbance of deeper functions of memory and thought on which human language is built.

The purpose of clinical interviews, then, is to use narrative language as a window to the schizophrenic mind. The motivation behind this dissertation is the idea that computational models should be able to do so as well. Consequently, this research represents an attempt to simulate a speaker with schizophrenia. A computational model of human narrative language is used to compare and distinguish a range of simulated illness mechanisms with respect to their ability to recreate the language-related symptoms of schizorphenia.

This chapter introduces the computational tools used to achieve this goal. First, DISCERN, a neural network-based model of human narrative language, is described. Based on the research literature, simulations of eight candidate illness mechanisms that could underlie schizophrenia are then discussed, and their implementation using the DISCERN model is described in detail. In the following chapters, the resulting candidate simulations of schizophrenia language are then evaluated experimentally.

3.1 The DISCERN Model

The original DISCERN model (Miikkulainen and Dyer 1991; Miikkulainen 1993) was the first integrated, neural network-based simulation of human story understanding and recall. It simulated human language behavior at several levels, from lexical access to overall discourse planning. DISCERN was able to understand, remember, and reproduce script-based stories and answer questions about them.

The original DISCERN was shown to exhibit important characteristics of human story processing, such as robustness to noise and realistic recall errors. However, it could only process stories that consisted of single script instances. In order to be useful as a model of schizophrenic language, it needed to understand and recall more realistic and complex stories. The model was therefore extended to process stories consisting of multiple scripts, using a new architecture to encode and retrieve episodic memories (Fidelman et al. 2005; Grasemann et al. 2007). Further extensions implemented in this dissertation include the ability to process emotional content of stories, and a filter for overly distorted language output. On the other hand, since the focus is on story recall, the question-answering modules of the original DISCERN were not included.

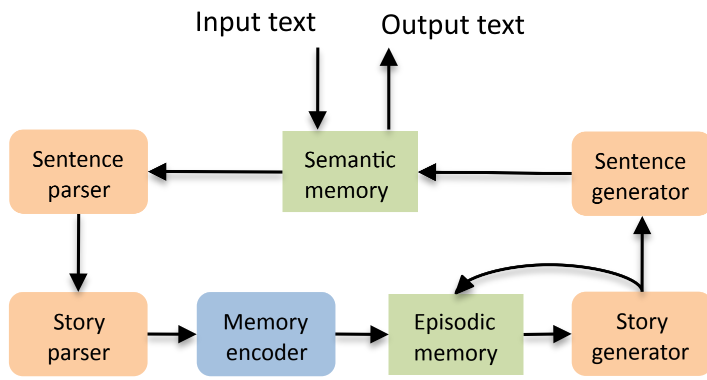


Figure 3.1: DISCERN is a neural network model of human story understanding and recall. The task of understanding and reproducing a story is achieved by a chain of modules, each building on the results of the previous module and providing input for the next. The figure shows the extended version of DISCERN used in this dissertation. It is able to understand and recall stories consisting of multiple scripts using the memory encoder module and a content-adressable episodic memory.

This section provides an overview of the current extended version of DISCERN, with special emphasis on the new memory architecture that enables the model to process multi-script stories. For simplicity, the extended model will be referred to as "DISCERN" in the remainder of this dissertation.

3.1.1 Architecture Overview

DISCERN reads and outputs natural language. Stories follow sequences of scripts, but are presented to the model as plain text, one word at a time. The task of understanding, remembering, and reproducing a story is achieved using a chain of neural network modules, each building on the results of the previous module in the series and providing input for the next (Figure 3.1). The modules communicate using patterns of neuron activations that encode word meanings. They are stored in a central lexicon, and are learned from input stories such that words that are used in similar ways have similar activation patterns. The other DISCERN modules are then trained in their tasks and learn to understand, remember, and paraphrase the stories.

For each story (such as the one in Figure 3.2) the word representations are presented to the sentence parser one at a time as a sequence of activation patterns. The sentence parser builds a representation of each sentence by concatenating the word representations that correspond to agent, predicate, indirect object, modifier, and direct object. At the end of each sentence, the sentence representation is passed on to the story parser. The story parser transforms sequences of sentences into script representations. A script representation consists of the name of the script and the words and emotion filling its slots. The sequence of script representations that constitute the story is stored in the episodic memory module in a compressed form, which is created by the memory encoding module using the Recursive Auto-Associative Memory (RAAM;Pollack 1990) architecture. To generate an output story, the story generator module translates the episodic memory representation into a sequence of sentences. Based on evidence of an editor function during human speech production (Fox Tree 2000), an output sentence filter evaluates and prunes this sequence. Finally the sentence generator module, last in the chain, reproduces the original word sequence for each sentence.

In this way, while DISCERN understands, remembers, and recalls the story, the content is tranformed from words to sentences, scripts, episodic memory traces, and eventually back to words. The remainder of this section describes each part of this process in detail, starting with the way in which the word representations in the lexicon are formed.

3.1.2 FGREP and the Lexicon

In the original DISCERN model (Miikkulainen 1993), the lexicon was an interconnected system of two self-organizing maps that translated between semantic and orthographic representations of the words used in the input stories. Since this part of the model is not a major

```
Emotion: Negative

[$job Vito Mafia head liked New-York famous gangster]
Vito was a gangster.                 [Vito was _ _ gangster]
Vito was the head of the Mafia.      [Vito was Mafia _ head]
Vito worked in New-York.             [Vito worked New-York _ _]
Vito liked his job.                  [Vito liked _ his job]
Vito was a famous gangster.          [Vito was _ famous gangster]

[$driving Vito _ scared airport LA recklessly _]
Vito wanted to go to LA.             [Vito wanted LA go _]
Vito entered his car.                [Vito entered _ his car]
Vito drove to the airport.           [Vito drove airport _ _]
Vito was scared.                     [Vito was _ _ scared]
Vito drove recklessly.               [Vito drove _ _ recklessly]

[$pulled-over Vito cop arrest(ed) _ murder _ _]
Vito was pulled-over by a cop.       [Vito was cop _ pulled-over]
The cop asked Vito for his license. [cop asked license his Vito]
Vito gave his license to The cop.    [Vito gave cop his license]
The cop checked the license.         [cop checked _ _ license]
The cop arrested Vito for murder.    [cop arrested murder _ Vito]

[$trial Vito _ walked cleared free murder good]
Vito was accused of murder.          [Vito was murder _ accused]
Vito was brought before the court.   [Vito was court _ brought]
Vito had a good lawyer.              [Vito had _ good lawyer]
The court cleared Vito of murder.    [court cleared murder _ Vito]
Vito walked free.                    [Vito walked _ free _]
```

Figure 3.2: An example input story for DISCERN about a gangster getting arrested for a crime committed in another story. The story consists of four scripts. The slot-filler representation of each script is on top, followed by the sentences of the script. Each sentence (left) is paired with its static case-role representation (right). During story understanding, DISCERN transforms such input stories from individual words to sentences, scripts, and finally episodic memory traces. Story recall later reverses this process to reproduce the individual words.

focus of the present work, a simplified version of the lexicon was used in order to reduce the overall complexity of the model. Instead of developing both orthographic and semantic representations of words, the current lexicon focuses on associating plain-text words with semantic word meanings, encoded as patterns of neuron activation. These patterns, called *word representations*, are fixed-length vectors of real numbers between 0 an 1. Similar words tend to have similar representations (in terms of Euclidian distance).

Apart from the words that encode semantic concepts, the lexicon also contains a special symbol " . " (period), which is used as an end-of-sentence marker, and is represented by a vector of zeroes. Additionally, it contains one pattern of random values for each script DISCERN learns to reproduce. The random patterns are used as script labels during story parsing and recall.

DISCERN's lexicon is accessed at three points during story recall. First, when DISCERN parses a story, the lexicon translates plain-text words into input activation patterns for the sentence parser. This translation is done by simple look-up: for each word, the associated representation is used as input to the sentence parser.

Second, when recalling a story, the ouput activation patterns produced by the sentence generator module are translated back into plain-text words by the lexicon. Since the sentence generator will generally not produce activations that precisely match a word representation in the lexicon, output word are selected by finding the nearest neighbor in Euclidean distance.

The third way in which DISCERN uses the lexicon occurs when the output filter is applied to the sentence representation produced by the story generator. The output filter (described in detail in Section 3.1.5) determines how well the words produced by the story generator match actual words in the lexicon, and prunes out the sentence from the output story if if the average similarity is below a threshold.

Note that in all cases, the way in which the lexicon is accessed is an abtraction of the original DISCERN lexicon implemented with a self-organizing map with plain-text labels:

When translating a word into an activation pattern, the weight vector of a neuron would be output based on its label. When looking up a plain-text word based on an activation pattern, the winner neuron would be decided based on Euclidean distance, and the corresponding label would be produced. The fundamental difference between the current lexicon and a self-organizing map lies in the way both are trained.

Words with similar semantics are represented by similar activation patterns in the lexicon. This property is achieved using the FGREP algorithm (Forming Global Representations with Extended backPropagation; Miikkulainen 1993), which develops word representations automatically based on the way words are used in the input text.

FGREP is most easily defined as a simple extension of standard backpropagation, where each input pattern is seen as an additional layer of weights that is adapted to the task at hand. Word representations are kept in the lexicon, and are used as input and target patterns in a traditional language processing network trained with backpropagation. The FGREP algorithm modifies the input patterns, which are then returned to the lexicon.

The major difference between FGREP and regular backpropagation is that the error signal is propagated one step further, from the hidden to the input layer:

$$\delta_{1i} = \sum_j \delta_{2j} w_{1ij},$$

where δ_{li} is the error signal of unit i in layer l; w_{lij} is the connection weight of unit i in layer l to unit j in the following layer. Note that this is a simple case of the normal backpropagation rule for error signal propagation where the activation function is the identity function. Using the error signal for the input layer, the input representations can then be modified:

$$p_{ci} = p_{ci} + \eta \delta_{1i},$$

where p_{ci} is the ith component of a word representation p_c, δ_{1i} is the error signal for input unit i, and η is the learning rate.

A convenient way to implement the FGREP algorithm is to add an additional input layer (layer 0) to the network, with one binary unit for each word in the lexicon. The new layer is fully connected to the original input layer. When a word is used as input, the corresponding binary unit's activation is set to 1, and all others to 0. The word representation p_c is then encoded by the connection weights: $p_{ci} = w_{0ci}$ for all i.

If the network is extended in this way, no modification of the backpropagation algorithm is necessary; the error signal need not be propagated further back than usual, and the rule for modifying the word representation becomes a special case of the normal rule

$$w_{lij} = w_{lij} + \eta\delta_{(l+1)j}o_{li},$$

where o_{li} is the output of unit j in layer l, which in this case would be 1 for the unit corresponding to the input word, and 0 for all others.

Intuitively, FGREP answers the question, "what input activation would have made the correct output more predictable?" The actual input of the network (the word representations) are then changed slightly in order to make the same sequence of words easier to predict in the future. If words are used in a consistent manner in the input sentences, changes to the word representations add up over time to reflect the way in which they are used. The resulting word representations reflect both grammatical role and semantic meaning of words, since co-occurrence of words is determined by both – e.g. nouns are often preceded by articles and followed by verbs; at the same time, the word "murder" is more likely to occur in a sentence that also contains the name of a gangster (as in the example story above). Both kinds of correlation will influence the representations developed by FGREP, since both enable the networks to predict their outputs more efficiently.

The only difference between word representations and connection weights in this algorithm is that word representations need to be used as output neuron activation patterns, and must therefore be constrained to the interval $[0, 1]$. This constraint is implemented in FRGEP by setting each component of the pattern to zero if it is negative, and to one if it is larger than one, after every FGREP iteration.

Importantly, in the resulting word representations, word meaning is distributed over the entire length of the pattern, not localized in specific neurons, as in the case of feature-based representations. This property of global rather than local semantic representations makes DISCERN more robust and allows it to generalize better: small activation errors lead to small (or non-existent) errors in word selection, and damage to part of an output pattern can be compensated by the remaining undamaged part.

3.1.3 Story Parsing

The task of parsing a story in DISCERN transforms an input story from a sequence of word representations to a (much more compact) sequence of script instances represented by their slot-filler representations. This transformation is achieved using the sentence parser and story parser modules (Figure 3.1), two simple recurrent networks (Elman 1990) that are trained to produce fixed-size representations of variable-length input sequences.

During story parsing, the sentence parser module receives a separate input pattern for each word in a story, plus the "period" pattern of all zeroes to mark the end of each sentence. While receiving the input words one at a time, the sentence parser builds a static output representation of the entire sentence. Sentences are represented by a list of five words, each corresponding to an input word that fills a specific case role in the input sentence. The case role structure used is generally [agent, predicate, indirect object, modifier, direct object]. For example, the sentence

```
Vince entered the LA airport
```

would be turned into the static case-role representation

```
[Vince entered _ LA airport],
```

where each plain-text word represents a pattern of neural activations, and the underscore (_) represents the blank pattern consisting of all zeroes. In this example, "Vince" is the subject, "entered" is the predicate, and "LA" modifies the direct object "aiport". There is no indirect object, so that slot is filled by the blank pattern.

In order to keep the representations compact while at the same time making a range of interesting sentences and constructions available to DISCERN, the individual case-role slots are overloaded slightly with related constructs. For example, the indirect object slot often contains prepositional objects, in which case the preposition used is implicit:

```
Vito drove to the airport. --> [Vito drove airport _ _]
```

Note also how articles are never part of the case role structure. The role of the modifier slot in particular is ambiguous and can contain adjectives or nouns that modify either the direct or the indirect object. It can also contain a possessive pronoun, as in

```
The cop asked Vito for his license.
--> [cop asked license his Vito].
```

In some cases, the modifier slot is also abused to introduce an infinitival phrase:

```
Vito wanted to go to LA.  -->  [Vito wanted LA go _]
```

As would be expected of a neural network architecture, DISCERN deals well with such underdefined sentence encodings. Prepositions and articles are inserted correctly, and creative and even inconsistent uses of case-roles is tolerated well, as long as they are the exception rather than the rule.

At the end of each sentence, the case-role representation produced by the sentence parser is passed on to the story parser module. Similar to the sentence parser, the story parser is a simple recurrent network that builds static representations of scripts from a variable-length sequence of sentences, much in the way the sentence parser builds sentence representations from a sequence of words. Scripts are represented as *slot-filler representations*: Each script is represented by a sequence of 8 words that encode the script's name and seven concepts filling its slots. For example, consider the third script in Figure 3.2:

```
Vito was pulled-over by a cop.        [Vito was cop _ pulled-over]
The cop asked Vito for his license.   [cop asked license his Vito]
Vito gave his license to The cop.     [Vito gave cop his license]
The cop checked the license.          [cop checked _ _ license]
The cop arrested Vito for murder.     [cop arrested murder _ Vito]
```

The sentence parser produces the sequence of case-role representations shown on the right; the story parser then encodes this sequence as an instance of the `$pulled-over` script:

```
[$pulled-over Vito cop arrested _ murder _ _]
```

The `$pulled-over` script has only four slots, so the other three contain the blank symbol. Note that in practice, slot-filler representations are concatenated activation patterns that each represent a word or script name in the lexicon. In this way, DISCERN's modules communicate using concepts and symbols encoded in the lexicon, but still retain the advantages of subsymbolic information processing.

In addition to the slot-filler representations, the story parser also builds a representation of the script's emotional valence. As mentioned earlier, each story in DISCERN has one of five emotion codes associated with it, ranging from very negative to very positive (–, -, +-, +, ++). These emotion codes are represented by word-sized activation patterns ranging from all zeroes for very negative (–) to all ones for very positive (++). The story parser learns to reproduce these patterns, which are then associated with each encoded script. The emotional valence always stays the same for each script of a story, so emotional information is an opportunity for DISCERN to resolve confusions between alternative ways to continue a story.

3.1.4 Memory Encoding

After DISCERN has read and decoded an entire story, the story's constituent scripts are represented by a sequence of slot-filler representations and their associated emotion codes. The task of the memory encoder module is to associate each script instance with a memory cue that allows stories in memory to be addressed and retrieved by content.

The memory encoder is a Recursive Auto-Associative Memory (RAAM; Pollack 1990), a neural network architecture that forms compact distributed representations of recursive data structures such as lists or trees. RAAM networks are feedforward networks trained to reproduce their own input, forcing them to form compressed representations of inputs in their hidden layer, which is smaller than the size of the input. These compressed representations can then be re-used as part of the input to the RAAM, and compressed representations of recursive data structures are formed as a result.

For example, consider a sequence of activation patterns $[a, b, c]$, each of size n. Say we would like to compress the entire list, and represent it as a single pattern of size h. This can be achieved using a RAAM network with h hidden units and $n + h$ input and output units. The first input to the network is a concatenation of pattern c and a special end-of-list pattern (usually n zeroes) denoted as $\Box$. After the network has propagated the input to the output layer, the hidden layer contains a compressed representation of the list $[c, \Box]$. It is useful to write this compressed list, LISP-like, as $(c\,\Box)$. The next input to the network consists of b and the compressed list $(c\,\Box)$ created in the previous step. The resulting activation in the hidden layer is the compressed list $(b(c\,\Box))$. The final input, a and the previous result $(b(c\,\Box))$ produce the compressed list $(a(b(c\,\Box)))$ in the RAAM's hidden layer (see Figure 3.3). Note that the same network could be used to compress a longer list, simply by continuing the process until all elements are included. Stepping backwards through a list and recursively re-using previous results in this way, RAAM networks can produce compressed representations of arbitrarily long lists. Also note that the final compressed list is not necessarily the only result of this process: compressed versions of partial lists, like $(b(c\,\Box))$, are produced as by-products along the way.

During the memory encoding process in DISCERN, the list of script representations produced by the story parser is compressed exactly in the way described above. More specifically, imagine that the three patterns $[a, b, c]$ in the previous example each represent one of the three scripts that make up a story. The memory encoder compresses this story by stepping backwards through the list of scripts. Each compressed partial list (and the final complete list) produced along the way is then used as a memory cue for the script used to create it, i.e. the compressed list $(c\,\Box)$ is the memory cue for script c, $(b(c\,\Box))$ is the cue for

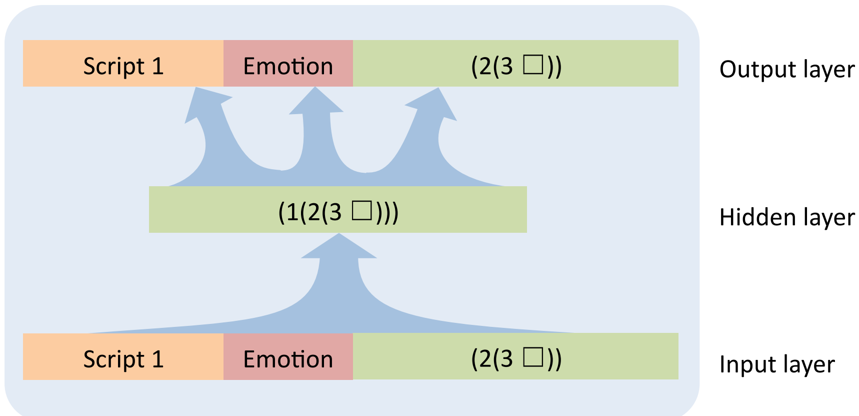


Figure 3.3: The memory encoder in DISCERN is a Recursive Auto-Associative Memory (RAAM; Pollack, 1990), a neural network that is trained to replicate its input at its output nodes, forcing it form a compressed representation of the input in its hidden layer. By reusing these hidden layer representations as part of the next input, RAAM can form fixed-size representations of recursive data structures like lists and trees. In DISCERN, compressed lists of scripts are created in this way and used as episodic memory cues.

script b, and $(a(b(c\,\square)))$ is the cue for script a. Figure 3.4 illustrates the memory encoding process in detail.

Each of the script-cue pairs produced by the memory encoder is called an *episodic memory trace*. In the original DISCERN, these traces were classified through a hierarchy of self-organizing maps, and a trace was created in the lateral connections of maps at the lowest level. Because episodic memory itself is not the main focus of the current work, its structure is abstracted in the current implementation: It is simply a store of script-cue pairs. During story recall, episodic memory is accessed using a memory cue; the memory trace that contains the most similar cue (in terms of Euclidean distance) is then retrieved.

One important feature of this memory mechanism is that the memory cues associated with each script are shaped by more than just the single script with which they are associated: Each cue encodes the entire remainder of a story at that point. This property implies that producing the correct sequence of memory cues needed to retrieve the scripts

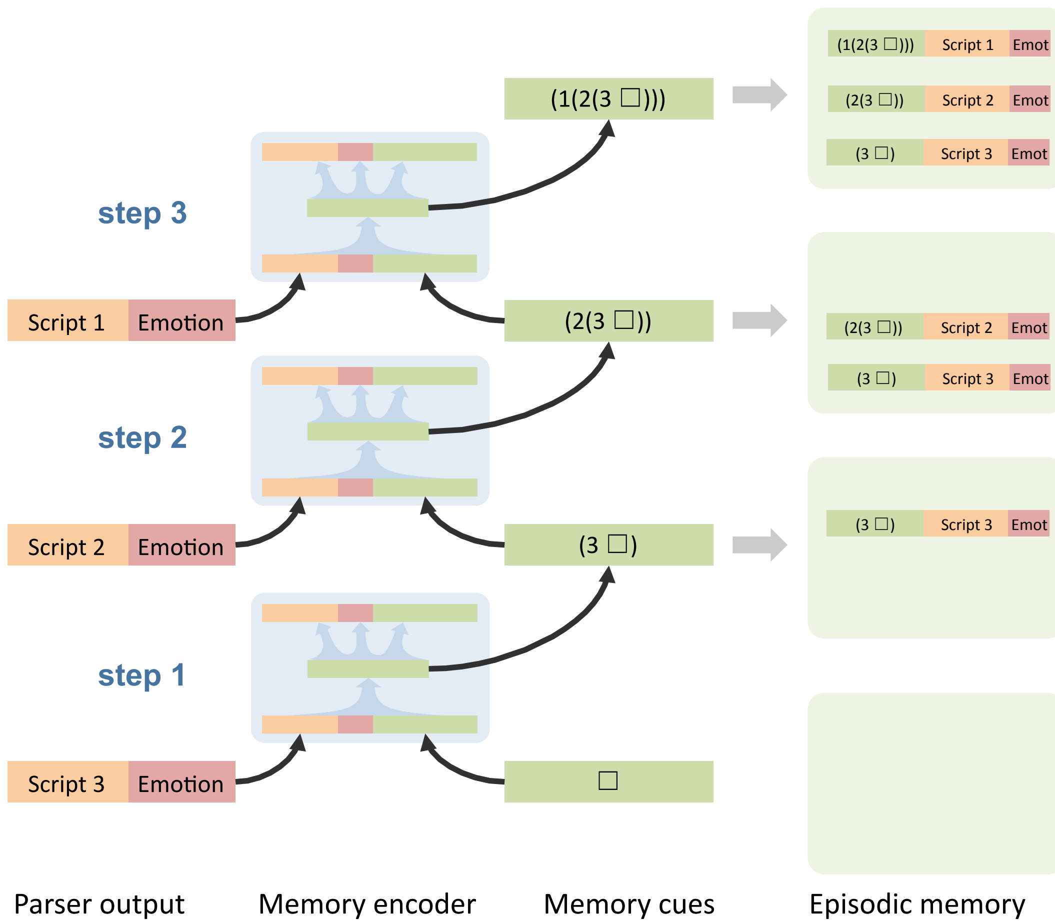


Figure 3.4: During the memory encoding process, each script of a story is paired with a memory cue, transforming the output of the story parser (left column) into content-addressable episodic memory traces (right column). Each script's memory cue is a compressed version of the remaining story, and represents DISCERNs discourse plan at that point – e.g. the cue for script 2 is the compressed version of scripts 2 and 3, denoteded (LISP-like) by $(2(3\square))$. The memory encoder builds these cues by stepping backwards (from bottom to top) through the scripts of a story, at each step creating a memory cue by combining a script with the memory cue produced previously. In this manner, stories of variable length can be compressed into a single distributed memory representation.

of a story in the right order would require maintaining a "discourse plan" at every point during story recall.

In DISCERN, RAAM representations of lists of scripts are used as cues to address episodic memory by content. Figure 3.3 shows a RAAM network that is being used to create a memory cue. Note how the network uses a cue (a compressed partial story) as part of the input to form the next cue in its hidden layer. In this way, the network steps backwards through a story, and produces a compressed representation of the rest of the story at each step, associating each new cue with the script used to create it. Figure 3.4 illustrates the encoding process.

3.1.5 Story Recall

Story recall is the process of transforming a story stored in episodic memory back into the original sequence of plain-text words. The story generator module is the heart of this process: It successively retrieves the memory traces that make up a story from episodic memory, and reproduces the story as a sequence of sentence representations.

Figure 3.5 shows the structure of the story generator module. It is a simple recurrent network that takes as its input a complete episodic memory trace (consisting of memory cue, slot-filler representation, and emotion code), and produces as its output a memory cue and a case-role representation of a sentence.

When recalling a story, the story generator is first cued with the first memory trace of that story. The initial output consists of the case-role representation of the story's first sentence, and a memory cue to episodic memory. Successive iterations of the story generator network produce a sentence and a memory cue each time. The memory cue is used to retrieve a memory trace from episodic memory, which is then used as the next input for the story generator.

Note that a memory trace encodes an entire script, which consists of multiple sentences. The story generator can produce only one sentence at a time, so the input memory

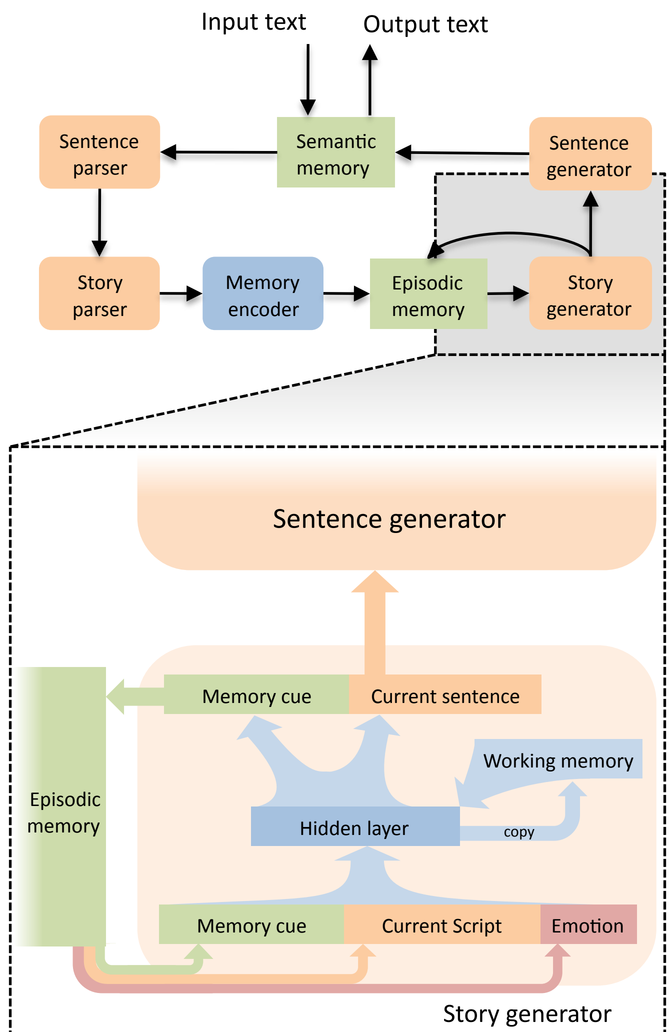


Figure 3.5: The story generator module in DISCERN is a simple recurrent neural network (Elman 1990). During story recall, the list of episodic memory traces encoding the current story are successively recalled from episodic memory. Each iteration of the story generator produces both a sentence representation and a memory cue. The sentence representation is passed on to the sentence generator; the memory cue is used to retrieve the next episodic memory, thereby determining the network's own next input. In this manner, the story generator can step through an arbitrary number of memory traces that encode a story, enabling it to process stories consisting of multiple scripts.

trace needs to stay the same until all sentences of the current script are produced. During recall of a single script, the story generator therefore produces the same memory cue repeatedly, while the output sentences change every time.

When the last sentence of a script is produced, the output memory cue changes, and the next memory trace (i.e. the next script) is retrieved from episodic memory. Figure 3.6 illustrates such a switch from one script to the next in detail.

Three snapshots of the story generator's input and output are shown during the switch from the second to the third script of the example story shown in Figure 3.2. The final memory cue, produced together with the last sentence of a story, is an end-of-story marker consisting of all zeroes. When this cue is passed on to episodic memory, an end-of-story memory trace is retrieved, and story recall ends.

Correct switching of memory cues during story recall is a complex task, since each script has a variable number of sentences, and each story has a variable number of scripts. Producing the correct sequence of cues therefore requires the story generator to maintain considerable internal state. The context layer of SRNs like the story generator is the only place where such networks can maintain such a state. The context layer therefore plays the role of working memory, and is therefore the target of several simulations of possible illness mechanisms.

Note also that, as mentioned earlier, the memory cues at each point during story recall are compressed versions of the entire remainder of the story. The ability to reproduce sequences of such complex structures for a large number of stories requires generalization. In other words DISCERN is forced to maintain a detailed discourse plan in the form of memory cues at all times during story recall.

Emotion codes, unlike memory cues, are encoded by very simple activation patterns, and are easy to reproduce. However, they are not unique to a story, so they cannot guide story recall by themselves. The emotional valence does stay the same throughout each story, however, so the emotion codes that form part of each memory trace hold valuable information that can be used to resolve ambiguity between alternative continuations of a story. Previous research has shown that DISCERN indeed makes use of the emotional information in this way (Fidelman et al. 2005).

Based on evidence of an editor function during human speech production (Fox Tree 2000), DISCERN includes an output sentence filter that estimates the distortion of every case-role representation produced by the story generator, and discards it if the distortion is

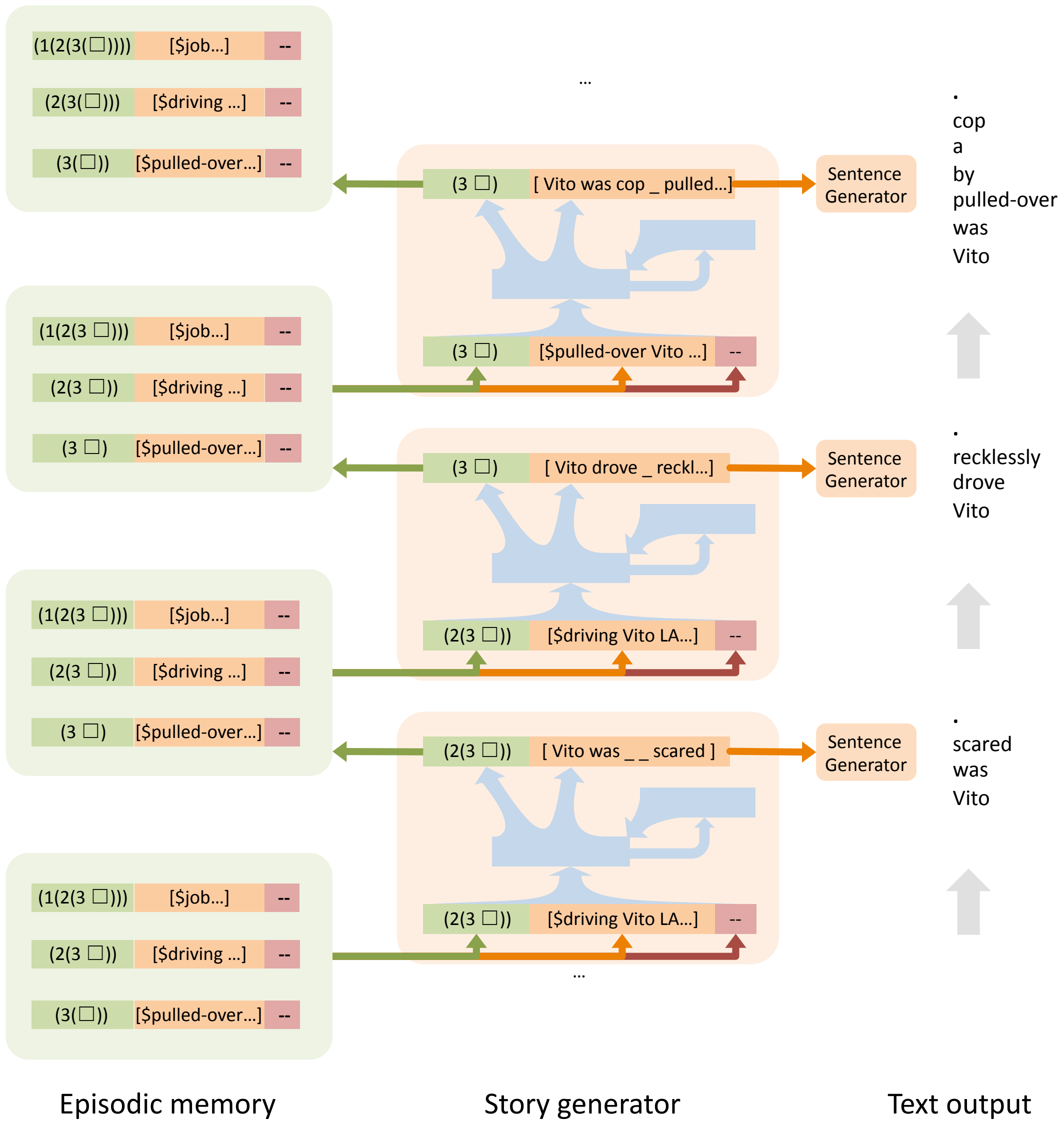


Figure 3.6: During story recall, the story generator steps through each sentence of the story, and accesses each memory trace encoding it. Three consecutive snapshots of the story generator's input and output are shown durng the switch from the second script ($driving) to the third ($pulled-over) script of the story shown in Figure 3.2. Time flows from bottom to top. *Bottom:* DISCERN reproduces the sentence "Vito is scared" in the second script of the story. The story generator produces a representation of the sentence, which is then passed on to the sentence generator (to the right). (continued on the next page)

(Figure 3.6 continued.) Additionally, it produces a memory cue that is used to retrieve the next input memory trace from episodic memory (on the left). In this case, the same memory trace as before is retrieved, since the script is not yet finished. *Middle:* DISCERN produces the last sentence of the script, “Vito drives recklessly.” The memory cue changes, and the memory trace for the third ($pulled-over) script is retrieved. *Top:* Using the retrieved memory trace, DISCERN now starts to reproduce the third script. By switching memory cues successively, the story generator can step through each script in the correct order. In this manner, scripts trigger subsequent scripts within a single story as is commonly done in symbolic script-processing systems (Schank 1999). In DISCERN, this model of narrative structure is given a subsmbolic connectionist implementation.

above a threshold. In this way, errors are reduced at the cost of reducing correct recall. The distortion $D(C)$ of a case-role representation C is estimated as the squared sum of the distortions of its constituent word representations:

$$D(C) = \sum_{i=1}^{5} |c_i - \hat{c}_i|^2, \tag{3.1}$$

where c_i is the ith component word of the case-role representation C, $\hat{c}_i$ is the closest match of c_i in the lexicon, and $|\cdot|$ denotes the Euclidean norm. This estimate relies on the fact that the lexicon is sparsely populated, so a distorted word representation is less likely to have a close match in the lexicon. Note that the only way to find the actual distortion of a word, rather than an estimate, would be to compare it to the correct target word, which is not known to the story generator during recall. In practice the estimate works well, especially in eliminating word errors that cross lexical categories. To make adjusting the behavior of the output filter more intuitive, the threshold is not set directly but is defined through a *filter strength* parameter s such that a sentence C is discarded when $D(C) > 1/s$. In this way, a higher filter strength means more output will be filtered.

The final step of the recall process is straightforward: Sentences that are not discarded by the output filter are passed on to the sentence generator, which is an SRN that takes a case-role representation of a sentence as its input and produces a sequence of individual word representations as its output. The input case-role representation stays the same over multiple calls, until the end-of-sentence pattern corresponding to “ . ” is produced. At this point, the current sentence is finished and processing of the next case-role representation can begin. The individual word representations are passed to the lexicon and converted into plain-text words, which form the final output of the DISCERN system.

3.2 Modeling Schizophrenia in DISCERN

DISCERN is a subsymbolic implementation of a symbolic theory. As such, it links two important conceptual levels, grounding human memory and language behavior in abstract neurons and synapses. As a basis for models of psychopathology, this is the core of DISCERN's strength: Underlying illness mechanisms can be simulated at the low level, and their manifestations observed in terms of high-level behavior.

The simulated illness mechanisms discussed in this section take full advantage of this opportunity. Eight illness models (or *lesions*) were implemented and evaluated in DISCERN. Each lesion simulates a low-level illness mechanism hypothesized to play an important role in causing symptoms of schizophrenia. Figure 3.7 gives a brief summary of the lesions, and indicates where in DISCERN they were applied.

The remainder of this section discusses the eight illness models in detail. The evidence supporting each illness model is reviewed briefly, summarizing the earlier discussion of possible illness mechanisms in Section 2.5. The implementation of each illness model in DISCERN is described, including the way in which the severity (or *lesions strength*) is adjusted in each case. Models of working memory impairment are introduced first, followed by semantic memory disturbances. Hyperlearning, a simulation of impaired dopamine transmission and memory consolidation, is described last.

1. **Working memory disconnection (WMD)**

   As discussed earlier in Section 2.5.1, loss of brain connectivity has been hypothesized to cause the symptoms of schizophrenia. Connections that have been found to be altered include both local cortical connections and connectivity between brain regions, and involve networks that are thought to be central to working memory function.

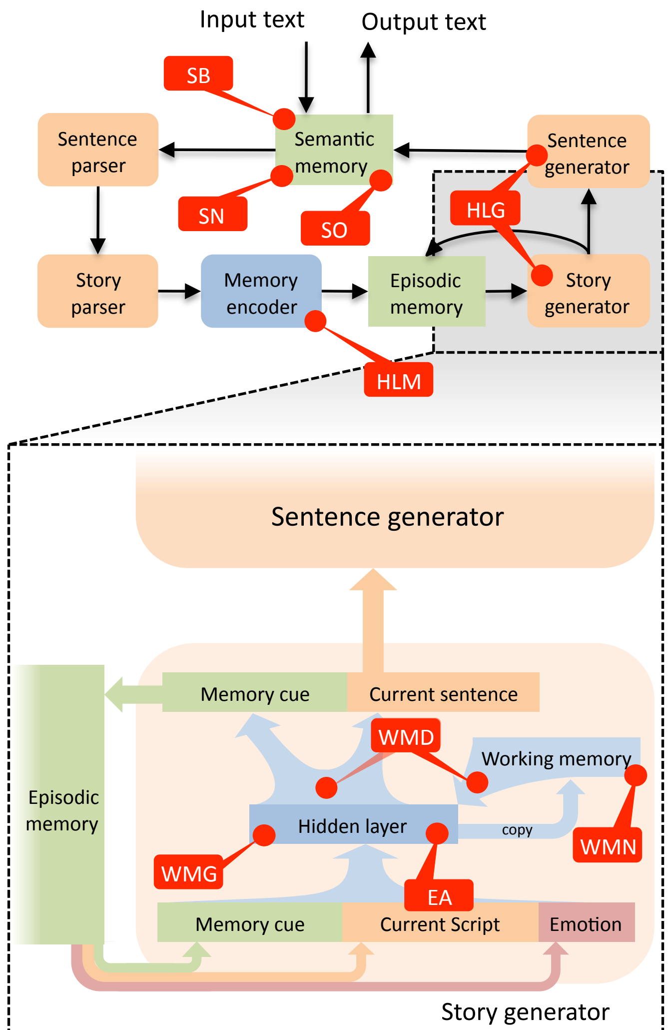


Figure 3.7: Overview of the illness models, or lesions, implemented in the DISCERN model. The labelled circles indicate to which part of DISCERN the lesion was applied. *WMD:* Working memory (WM) disconnection (pruning WM → hidden connections, and, in the extended model, hidden → output connections as well); *WMN:* WM noise (added to the activation of WM neurons); *WMG:* WM gain reduction (less steep sigmoid in WM neurons); *EA:* Excessive arousal states (varying bias weights of story generator WM neurons); *SN:* Semantic memory distortion (adding noise to semantic word representations); *SO* Semantic overactivation (increasing activation of semantic word representations), *SB:* Semantic blurring or overpriming (blurring semantic word representations); *HLM, HLG:* Hyperlearning (amplified backpropagation learning, applied to the memory encoder (HLM) or the story and sentence generators (HLG). In the three-parameter hyperlearning model (Chapter 5), both HLM and HLG were used at the same time. The different illness models will later be evaluated as simulations of schizophrenic language.

The primary working memory component that governs language production in DISCERN is part of the story generator network. Recall that the story generator is a simple recurrent network, i.e. each time the network is activated, it stores its hidden-layer activation in its context layer. The context layer in turn has a full set of trainable forward connections to the hidden layer, so the network has access to its own previous internal state. This hidden-context-hidden layer loop effectively enables the network to learn, represent, and act on patterns in time, such as scripts, which are patterns of propositions in time.

The WMD lesion therefore targeted this component of the story generator network. In order to simulate altered connectivity, DISCERN's working memory was partially disabled by pruning the forward connections from the context layer to the hidden layer of the story generator network. The parameter that determined the severity, or strength, of the lesion was the pruning threshold: All connections whose absolute weight fell below the threshold were cut, i.e. set to zero.

Note that, even though working memory circuits are implicated in the literature on functional disconnection in schizophrenia, disconnection does not necessarily have to involve only working memory. The same pruning method can be used to partially disable other comonents of the network. For example, in experiment I (Chapter 5), both the working memory connections and the hidden to output layer connections in the story generator were pruned independently.

2. **Working memory noise (WMN)**

   Working memory dysfunction is strongly associated with schizophrenia (Aleman et al. 1999), and may have other causes than altered brain connectivity. As discussed in section 2.5.2, abnormal DA activity may cause noisy and otherwise impaired information processing in WM networks.

In DISCERN, noisy information processing of this type was modeled using the working memory of the story generator network. Every time hidden activations were copied to the context layer, Gaussian noise with zero mean was added to the activations. The strength of the lesion was controlled by changing the variance of the noise; higher variance meant higher levels of noise distortion.

3. **Working memory gain reduction (WMG)**

   Changing the gain of the activation function of artificial neurons has been investigated previously as a simulation of changed dopaminergic neuromodulation in schizophrenia (Cohen and Servan-Schreiber 1992; Sections 2.5.2 and 2.6). Since DISCERN's working memory, discourse planning, and context processing during story recall all happen in the story generator, this module was again the target of this lesion. Working memory gain reduction (WMG) was implemented by changing the slope of the sigmoid activation function of the hidden-layer neurons in the story generator. By default, all activation functions in DISCERN are sigmoids with unit slope, and are not adapted during training. The WMG lesion reduced the slope; the lesion strength parameter controlled the amount by which it was reduced.

4. **Excessive Arousal (EA)**

   Historically, a popular view has been that symptoms of schizophrenia, and psychosis in particular, are caused by a state of constant cortical overarousal, or by susceptibility to such a state (Kornetsky and Eliasson 1969; Schlör et al. 1985). A recent study on the genetics of catecholamine function Arnsten (2007) suggests how such a susceptibility could occur: Both schizophrenia and bipolar disorder are associated with genetic changes that "may render patients vulnerable to profound stress-induced PFC dysfunction including symptoms of thought disorder." Furthermore, one of the earliest neural network-based models of schizophrenia was based on the theory of overarousal (Grossberg and Pepe 1970; see Section 2.6).

Elevated arousal at a neuronal level was simulated in DISCERN by increasing the bias of all WM neurons in the story generator by a fixed amount. The strength of this lesion was adjusted by changing the value added to the bias.

5. **Semantic blurring (SB)**

   The *hyper-priming* hypothesis (Section 2.5.3) suggests that excessive spreading activation in semantic maps is a major contributor to the symptoms of thought disorder in schizophrenia, and has been the subject of previous computational models (Spitzer 1997; Leeson et al. 2005; see Section 2.6 for details).

   In DISCERN, hyper-priming, i.e. excessive coactivation of related words in the lexicon was simulated by *semantic blurring*, i.e. instead of a word representation w the lexicon produced a mixture of w and its neighbors. These "blurred" word representations were calculated in the following way: for each word w, the lexicon was sorted by Euclidean distance to w. The blurred version of w is then defined in the following way (note the convex sum for $0 < \beta < 1$):

$$\mathrm{blur}(w) = \sum_{i=0}^{N-1} (1-\beta)\beta^i w_i$$

   where w_i is the ith closest word to w, and $w_0 = w$. N is the size of the lexicon, and $0 < \beta < 1$ is a "blur factor" – larger β means more blurring. The lesion strength for the SB lesion is adjusted through β.

6. **Semantic memory noise (SN)**

   Altered verbal fluency and object comparisons suggest that semantic memory itself, not just access to it, may be disorganized in TD schizoprenic patients (Goldberg et al. 1998; Tallent et al. 2001; Section 2.5.3). In order to simulate distortion or disorganization of semantic memory, Gaussian noise with zero mean was added to the word representations in DISCERN's lexicon. Note that this lesion is intended to model changes to the semantic memory itself, so noise was added only once, not continuously during lexical access. The severity, or strength, of the semantic distortion was controlled by adjusting the variance of the noise.

7. **Semantic memory overcativation (SO)**

   Several recent functional imaging studies reported cortical overactivation in regions involved in semantic memory and language processing during tasks that involved semantic association and lexical access Kuperberg et al. 2007; Assaf et al. 2006. The SO lesion in DISCERN consequently simulated overactivation of semantic memory networks by adding a constant bias to word representations produced by the lexicon. The strength of the SO lesion was regulated by adjusting this bias.

8. **Hyperlearning (HLM, HLG)**

   Hyperlearning is a version of the DA hypothesis based on Kapur's (2003) theory that increased midbrain DA release leads to abnormally enhanced motivational salience, which in turn causes psychotic symptoms (see Section 2.5.2 for details). Hyperlearning extends and formalizes this theory by proposing a concrete mechanism by which this could occur: Aberrant salience of experience leads to overly intense memory consolidation. This hypothetical "hyperlearning state" was simulated in DISCERN by performing additional backpropagation learning at elevated learning rates. Hyperlearning was always applied for 500 additional epochs (i.e. iterations of the entire story corpus) after DISCERN was fully trained.

   Hyperlearning was applied to either the memory encoder network (HLM) or the generator modules (story generator and sentence generator; HLG), or both. In the case of HLM, the down-stream generator modules were also trained (using normal learning rates) in order to enable them to compensate for changing memory cues.

Not all hypotheses about possible illness mechanisms could be simulated meaningfully in DISCERN. For example, loss of attentional control and impaired theory of mind are currently not accessible to the model (but see Chapter 7). Nevertheless, the illness models span a wide range of current theories, and reflect the emphasis that is currently placed on the role of impaired semantic memory and disturbances involving dopamine and/or working memory.

3.3 Conclusion

The purpose of this chapter was to introduce the computational modeling tools that will be used in this dissertation to create and evaluate simulations of schizophrenic language. DISCERN, a connectionist model of human story understanding and recall, was descibed in detail. Recent extensions of the model were discussed, including the ability to process multi-script stories, emotions, and a filter mechanism that reduces errors at the cost of lower overall language output.

Based on the DISCERN model, eight simulations of candidate illness mechanisms that could underlie symptoms of schizophrenia were then introduced. Each of these illness models was motivated by a specific hypothesis about illness mechanisms in schizophrenia. The implementation in terms of the DISCERN model was described in each case.

DISCERN and the different illness models will ultimately be evaluated as simulations of schizophrenic language. However, in order to do so, a simulation of normal human language based on DISCERN must be created first as a starting point. The development of this "healthy" DISCERN model is the topic of the next chapter.

Chapter 4

Developing a Healthy Model

DISCERN is a complex model, where multiple functional layers combine, learn, and perform together to create the final language behavior. Much of its power as a model of schizophrenic language derives from the complexity of the underlying healthy language: When story processing breaks down due to simulated illness mechanisms, it may do so in equally complex and unexpected ways. In order to create the most informative models of schizophrenic language possible, it was therefore necessary to push the limits of the DISCERN system and develop concrete, running instances of the model whose language behavior was as rich and varied as possible. The purpose of this chapter is to describe the steps that were necessary to achieve this goal.

First, the model was implemented and integrated into the existing parallel computing infrastructure. Second, a corpus of input stories was created that was large and complex enough to make it possible to observe and quantify subtle changes in recall and language performance. Based on the vocabulary used in these stories, a lexicon of word representations was then developed using the FGREP algorithm described in section 3.1.2. Finally, training methods and schedules were developed that allowed DISCERN to learn the entire story corpus almost perfectly. The resulting "healthy" instances of the DISCERN model form the basis of all computational experiments discussed in this dissertation.

4.1 Implementation Details

The C implementation of the extended DISCERN model was based on the *proc* module that was part of the original DISCERN implementation by Risto Miikkulainen (available at http://nn.cs.utexas.edu/?discern), and on previous code extensions by Hong Ming Yeh and Peggy Fidelman (Fidelman et al. 2005). The original code was extensively redesigned and optimized for modern computer architectures.

Comparisons between alternative lesions, especially for experiment I (Chapter 5), made it necessary to explore the parameter space of different lesions extensively. In order to make this possible, tools were developed to integrate the model into the existing parallel computing infrastructure. All computational experiments were run on the Mastodon cluster (http://www.cs.utexas.edu/facilities/accommodations/condor/mastodon) using the condor job-scheduling system.

The findings reported in the following chapters are based on millions of different combinations of lesion and filter parameters tested. For each of these combinations, detailed statistics on the language performance of 30 DISCERN systems were collected. Specialized tools were created, mostly using Python (www.python.org van Rossum and de Boer 1991)), to manage the large amount of experimental data created in this way. Data visualization, including all plots in this dissertation, was done using matplotlib (Hunter 2007,; http://www.matplotlib.sourceforge.net). Additionally, a website that allows users to explore the data for experiment II (chapter 5) was developed using Python and gnuplot, both as a demo and to help communication within the project.

4.2 Story Corpus

In the original DISCERN model, all stories consisted of a single script, which limited DISCERN to learning stories describing a single, relatively uncomplicated event, like eating in a restaurant or getting on a plane. The current extended model is able to store and recall

```
Emotion: positive

[$drink I Stacy met Moe's-Tavern table wine _]
I went to Moe's-Tavern          [I went Moe's-Tavern _ _]
I sat at a table                [I sat table _ _]
I ordered wine                  [I ordered _ _ wine]
I drank the wine                [I drank _ _ wine]
I met Stacy at Moe's-Tavern     [I met Moe's-Tavern _ Stacy]

[$relation I Stacy liked trusted _ my girlfriend]
Stacy was my girlfriend         [Stacy was _ my girlfriend]
I liked Stacy                   [I liked _ _ Stacy]
I trusted Stacy                 [I trusted _ _ Stacy]

[$person Stacy 20s ponytail New-York movies books compact]
Stacy was in her 20s            [Stacy was 20s her _]
Stacy had a ponytail            [Stacy had _ _ ponytail]
Stacy was from New-York         [Stacy was New-York _ _]
Stacy drove a compact car       [Stacy drove _ compact car]
Stacy liked movies              [Stacy liked _ _ movies]
Stacy liked books               [Stacy liked _ _ books]

[$talking I Stacy  liked liked kiss books long]
I talked to Stacy about books   [I talked Stacy about books]
I liked books                   [I liked _ _ books]
I talked to Stacy a long time   [I talked Stacy long time]
I liked talking to Stacy        [I liked Stacy _ talking]
I gave a kiss good-bye to Stacy [I gave Stacy kiss good-bye]
```

Figure 4.1: An uneventful example story from the personal context: " I " meet my girlfriend Stacy for a drink, and we have a conversation about books. Most stories in the personal context are, like this one, told from the first-person point of view. The " I ", or self, is overrepresented in the story corpus in order to simulate the concept of the person experiencing and telling the stories.

stories that can in principle be composed of an unlimited number of scripts. In practice, DISCERN's capacity to process multi-script stories (and mine to come up with them) is limited to about seven scripts. Nevertheless, this ability makes a wide range of more complex stories accessible. Scripts can be combined and recombined in new ways, or can be repeated within a story, expressing different content by using different slot-fillers. Stories can also share multiple of scripts, establishing stereotypical, repeating sequences of scripts that can themselves be viewed as scripts that are implicitly learned by the model. Overall, multi-script stories make DISCERN a more useful model by creating a more difficult task, demanding more complex behavior, and giving DISCERN more opportunities to fail in interesting ways.

The story corpus designed for this dissertation is an attempt to make use of these opportunities, and to stretch as far as possible the kind and number of stories that DISCERN can learn. It was also designed with the goal of this dissertation in mind: The stories need to make it possible to investigate the kinds of language disturbance observed in patients with schizophrenia. Specifically, failures of context, of continuity, and of character slotting need to be observable during story recall. Opportunities for content to intrude from one story into another had to be created. The remainder of this section describes the corpus in detail, and addresses how these issues were addressed.

The corpus contains 28 stories. Each one is a sequence of three to seven scripts, and contains between nine and 35 sentences. All stories taken together contain 550 single sentences in 120 script instances. The size of the vocabulary is about 160 words, including 20 names or descriptions of story characters (e.g. "Frank" or "lawyer"). The entire story corpus is reproduced in Appendix 8, including script and sentence representations.

The stories in the corpus are divided into two groups, defining two distinct categories, or *contexts*. The first (the "personal context") consists of stories from the life of a character called "I" (referred to as the *self* in the following). The stories in the personal context mostly described his relationships and experiences, and attempted to create a somewhat coherent slice of his life. Most of these stories are told from the first-person point of view of the self, who is overrepresented in this way in order to simulate the central role of the self in autobiographical human memory. An example of such a story is shown in Figure 4.1. The stories in the personal context span the whole range of emotional tone, from very negative to very positive.

The second group (the "gangster context") contains stories with a very different theme, and mostly negative emotional tone. They concern a group of Mafia-type gangsters who mostly engage in stereotypical gangster activities, committing crimes, killing each other, and occasionally getting caught by the police. The stories in this context are intended to simulate the impersonal stories to which humans are exposed, e.g. through movies or the news. In contrast to most personal stories, they are told from the third-person point of view of varying characters; the self character does not appear. An example of a gangster story was previously shown in Figure 3.2.

Note that the two story contexts are entirely implicit: DISCERN is given no direct information that would make it possible to distinguish one context from the other, or even to decide whether or not stories can usefully be divided up in this way. At the same time, contextual cues are everywhere in the story content. For example, words that are unique to one context like `gun` or `wedding` make it easy in principle to decide what the current context is. This contextual information can then help story understanding and recall, e.g. by resolving ambiguity. The data reported below on the errors that occur in heathy systems suggest that DISCERN indeed learns to use these contextual cues.

Some of the most important and unambiguous contextual cues are names of story characters: All ten named characters are context-specific, i.e. each appears in the personal context or the gangster context, but not both. Five characters (three gangsters and two policemen) are unique to gangster stories, and five (including the self) appear only in personal stories. Unnamed characters that appear in the stories, like `cop` or `girlfriend`, are not always unique to a context.

In order to encourage DISCERN to use contextual information, and also to make failures of context easier to observe in the output language, there are several pairs of characters that are similar, but belong to different contexts. For example, the self's boss `Joe` is in many ways similar to the Mafia boss `Vito`. At any point in a story, the easiest way to decide between the two would be to consider the current context. On the other hand, if

Table 4.1: Overview of the 14 scripts used to build the story corpus. The length of each script is given in sentences.

Script Name	Context	Content	Length
$person	Both	Description of a story character	6
$job	Both	Desciption of a character's job	5
$relation	Both	Relationship between two persons (e.g. boss, finacee)	3
$flight	Both	Getting on a plane and flying somewhere	7
$drink	Both	Having a drink and meeting someone	5
$driving	Both	Driving somewhere in a car	5
$drunk	Both	Getting drunk	3
$pulled-over	Both	Getting pulled over by the police	5
$trial	Both	Being accused of a crime in court	5
$talking	Both	A conversation between two characters	5
$plan	Both	Two characters plan an occasion	4
$occasion	Both	Generic occasion, including wedding and bombing	5
$being-after	Gangster	An organization, e.g. the police, is after someone	5
$investigation	Gangster	The police investigates a crime	4

DISCERN learns to rely on such information, and context processing is later impaired, this reliance on contextual cues should result in an observable tendency to confuse characters across contexts.

All stories were assembled from 14 different scripts, briefly summarized in table 4.1. Examples of each script and its slots can be found in Appendix 8. Most scripts describe stereotypical sequences of events such as meeting someone for a drink or being pulled over by the police. Additionally, several scripts (`$person`, `$job`, and `$relation`) do not describe an event, but instead describe a person or a relationship between two persons. The purpose of these scipts is to create opportunities for DISCERN and the FGREP algorithm to develop detailed and complex representations of the agents in the story corpus.

Most scripts were designed to be used in both personal and gangster contexts, and in stories with either positive or negative emotion. For example, there is no specific script describing a crime — instead, there is a general `$occasion` script that can encode various very different events, depending on the concepts used to fill the slots. In story #26, for example, the `$occasion` script is used to describe how `Vince`, a Mafia hitman, kills `Tony`, another gangster:

```
[$occasion Vince Vince phone-call killed Starbucks Tony murder]
Vince entered Starbucks for murder.     [Vince entered ...]
Vince killed Tony.                      [Vince killed _ ...]
The murder was a success.               [murder was _ _ ...]
Vince made a phone-call.                [Vince made _ _ ...]
Vince smoked a cigarette.               [Vince smoked _ ...]
```

In story #17, the same script is used to describe a harmless wedding:

```
[$occasion I Mary speech kissed Four-Seasons Joe wedding]
I entered the Four-Seasons for wedding. [I entered ...]
Mary kissed Joe.                        [Mary kissed ...]
The wedding was a success.              [wedding was ...]
I gave a speech.                        [I gave _ _  ...]
I drank champagne.                      [I drank _ _ ...]
```

In this way, shared scripts between contexts create opportunities for DISCERN to cross over between contexts, and again encourage the use of context. In the same way, coincidental shared structure between real-life and imaginary stories in humans may create opportunities for derailments or for the formation of delusional ideas. In the instances of the `$occasion` script above, note also that the script content is mostly, but not entirely, determined by inserting slot fillers into otherwise fixed sentence structures. For example, the content of just one slot (`speech` vs. `phone-call`) determines the structure of an entire sentence (`I gave a speech.` vs. `Vince made a phone-call.`). DISCERN can make use of this kind of flexible script definition, as long as most scripts follow a regular structure.

4.3 Developing a Lexicon

The story corpus described above containes approximately 160 unique words, including 78 nouns, 44 verbs, 18 adjectives or adverbs, and 10 prepositions. Apart from the prepositions, other closed-class words are definite and indefinite articles, and the pronouns "his", "her", "my", "me", and "that". Of the nouns, 20 described story characters or groups of people like "police" or "Mafia". Ten of the words describing characters were names ("I" is considered a name here rather than a pronoun); the rest were more abstract descriptions of characters like "cop" or "fiancee".

Before DISCERN's processing modules could be trained to process the story corpus, it was necessary to develop meaningful word representations for these words based on the way in which they were used in the input stories. This was achieved using the FGREP algorithm described in Section 3.1.2.

Since both the story corpus and the vocabulary were many times larger than any that had been used in DISCERN previously, some preliminary experiments were necessary to determine the behavior of the algorithm. These experiments led to several observations that proved to be useful in developing a large lexicon with FGREP. First, the success of FGREP depended to some degree on the size of the word representations, but above a certain size, larger representations lead to no additional benefit. In this case, representations of size 12 turned out to be close to optimal – the default size that was also used in the original DISCERN model.

Second, FGREP worked well when word representations were trained while sentences were processed by DISCERN's sentence parser module. Using any other combination of modules reduced the quality of the word representations, i.e. the similarity of representations did not reflect word similarity as well. One reason for this difference in FGREP performance between modules could be that the input to the sentence parser network is a single word representation at a time. The input layers of DISCERN's other processing modules all consist of multiple words, which may lead to a noisier, less word-specific error signal. It is also possible that the patterns of use and co-ocurrence in sentences are clearer and more useful than, say, in slot-filler representations.

The quality of word representations also depended significantly on the duration of training, on the learning rates for connection weights and for word representations, on network size, and on a range of other parameters. The best word representations seemed to result when relatively small sentence parser networks were trained for a relatively short period, and when the learning rate for word representations was significantly higher than the one used for network connections. Intuitively, the opposite should be true for each of these parameters: longer training and larger networks should make learning easier; learning rates should be higher for the network connections in order to avoid moving-target effects (Miikkulainen 1993). The reason for this behavior remains to be investigated in future work.

The FGREP algorithm did, however, produce a set of high-quality word representations. The best results were trained using a sentence parser network with 12 input units, 150 hidden units, and 60 output units. FGREP training lasted for 500 epochs, i.e. the network was exposed to each sentence in the corpus 500 times. The learning rate was 0.01 for network connections and 0.1 for word representations. The values of the word representations were clipped to the interval $[0, 1]$.

By the end of training, the word representations reflected the similarities between the concepts well: words whose representations were close tended to denote similar concepts, and usually belonged to the same lexical category. Table 4.2 illustrates this tendency by listing the four closest words (by Euclidean distance) for a representative subset of words in the lexicon. Note, for example, how the names of story characters form a tight and well-defined cluster. With only a single exception, the words closest to each name are other names. Note also that the closest name to `Joe` is `Vito`, reflecting the similarity of the self's boss and the Mafia boss mentioned in the previous section. Other word categories form similar clusters, although they are generally not as tight as the cluster of story characters.

Table 4.2: The four words with the most similar FGREP representations (by Euclidean distance) are shown for a representative subset of the lexicon. Similarity of word representations corresponds well with similarity of word meaning.

Word	Closest	Second	Third	Fourth	Distance
I	Vince	Tony	Joe	Fred	(0.23–0.43)
Stacy	Kate	Mary	Fred	Vito	(0.08–0.31)
Mary	Stacy	Kate	Vito	Bob	(0.15–0.28)
Joe	Vito	Mary	Kate	Fred	(0.21–0.32)
Tony	Vince	I	the	Joe	(0.22–0.48)
Vince	Tony	I	Fred	Joe	(0.22–0.41)
Vito	Joe	Mary	Stacy	Bob	(0.21–0.31)
man	Joe	Tony	I	Vito	(0.47–0.60)
boss	girlfriend	fiancee	pulled-over	mother	(0.41–0.83)
co-worker	mother	friend	fiancee	brought	(0.67–0.90)
girlfriend	boss	mother	pulled-over	fiancee	(0.41–0.78)
friend	mother	co-worker	girlfriend	brought	(0.45–0.99)
mother	friend	girlfriend	co-worker	accused	(0.45–0.80)
doctor	gangster	beer	cigarette	late	(0.32–0.91)
Mafia	police	job	bag	St-vincent's	(0.53–0.83)
police	Mafia	job	bad	Moe's-tavern	(0.53–1.02)
New-york	Starbucks	LA	city-hall	St-vincent's	(0.50–0.72)
Chicago	LA	St-vincent's	New-york	fine	(0.22–0.79)
airport	Four-seasons	city-hall	guns	books	(0.57–0.66)
Starbucks	city-hall	New-york	Four-seasons	St-vincent's	(0.47–0.63)
city-hall	Starbucks	airport	Four-seasons	to	(0.47–0.65)
wedding	murder	meeting	bombing	Vince	(0.39–0.60)
bombing	murder	meeting	the	wedding	(0.39–0.54)
hated	feared	distrusted	trusted	$trial	(0.72–0.80)
trusted	loved	feared	distrusted	kissed	(0.11–0.52)
loved	trusted	distrusted	feared	kissed	(0.11–0.60)
liked	drank	feared	$plan	Kate	(0.90–1.00)

Table 4.2 also illustrates another interesting feature of the FGREP representations: the lists of similar words reflect a mixture of similarity in grammatical role and similarity in semantic meaning. This effects is most clearly visible in the cluster of character names: their grammatical roles are virtually identical, but the internal structure of the cluster nevertheless reflects more subtle similarities, e.g. `Mary` is closest to `Stacy`, and `Tony` is closest to `Vince`. In the same way, the positive verbs `loved` and `trusted` are closer to each other than to negative verbs like `feared`. This observation shows one of the main strengths of the FGREP approach: When learning word representations from actual language, lexical categories provide the strongest and most immediate organizing principle, but once words are ordered according to that principle, more subtle differences in the way words are used make it possible to capture semantic meaning as well. Figure 4.2 shows a principal component analysis of the 102 words that were used as slot fillers in the story corpus. Words cluster relatively well according to grammatical role and semantic categories. Again, several levels of more or less fine-grained neighborhood relationships expressing different levels of similarity are visible. Note, for example, the cluster containing "rusty", "nice", and "compact", three words exclusively used in the story corpus to describe cars. Also note how the internal structure of the tight cluster of character names. The two components shown account for 43% of the data variance.

In summary, the FGREP algorithm was able to learn word representations that reflect similarity of both lexical category and semantic meaning well. Both of these similarities were learned exclusively from the way words were used in the input stories. Meaningful word representations such as these are essential because they ensure that small errors in network output translate into equally small errors in word selection, making DISCERN more robust and enables it to degrade gracefully with damage.

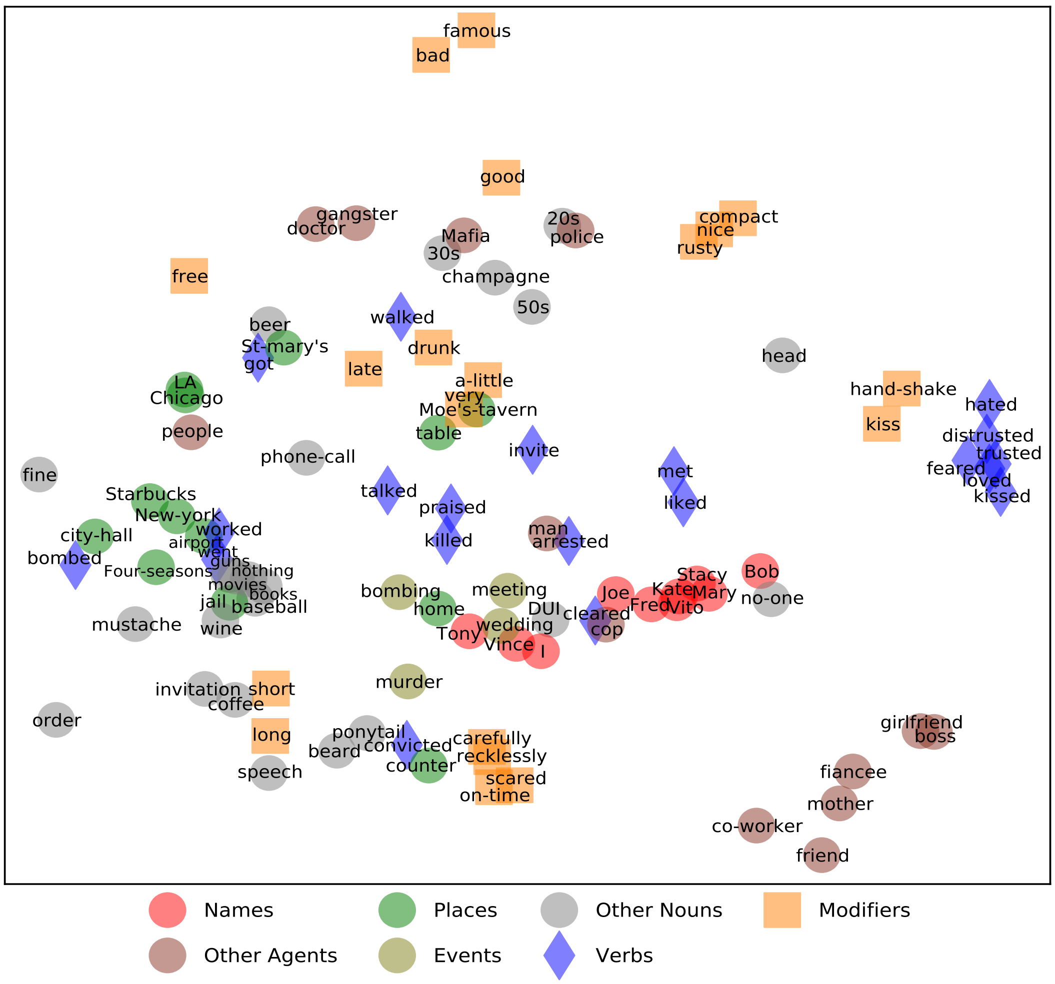


Figure 4.2: Principal Component Analysis of word representations learned by the FGREP algorithm. The set of 102 words that were used as slot fillers in the story corpus are shown. Clusters form at different levels of detail, according to a mixture of grammatical role and semantic similarity. Note e.g. the fine-grained structure in the central cluster of character names. The two principal components shown account for 43% of the data variance. Some labels (but not markers) were moved slightly to avoid overlap.

4.4 Network Training

With the corpus and the word representations in place, the next step was to develop effective methods to train the DISCERN model to understand and reproduce the stories. Network sizes and parameters were determined empirically. Sentence parsers and generators had 250 hidden neurons, story parsers had 225, and story generators had 150. Memory encoder networks had 48 hidden neurons.

Modules were trained in a chain, with the output from one module used as the input for the next. Starting with the sentence parser module, new modules were added to the chain successively as meaningful input became available during the course of learning.

The learning rate for each module was set to 0.4 times the average output error of the module (root mean squared error, averaged over all outputs of the network during the previous training epoch). In this way, the learning rate for each module decreased automatically as the output error decreased during training, which allowed for fast weight changes during early training as well as fine-tuning of network response at low learning rates later on.

It should be noted that the usefulness of this kind of adaptive learning rates in backpropagation training is controversial (Sarle 1997; Bertsekas and Tsitsiklis 1996). Nevertheless, it works well in the case of DISCERN. One particular advantage is that learning rates adapt independently for each module, which reduces the complexity of training schedules considerably.

A total of 70,000 backpropagation learning epochs were employed overall for each DISCERN system, even though only a subset of all modules were trained during most epochs. The following schedule was used to determine which modules to train at each point during training:

1. **Until epoch 30,000:** Train only the sentence parser module.

2. **Epoch 30,000 – 40,000:** Train the story parser module. The sentence parser module is still running to provide input, but is not being trained.

3. **Epoch 40,000 – 50,000:** Train the memory encoder and story generator modules. The sentence and story parsers are running to provide input, but are not being trained.

4. **Epoch 50,000 – 60,000:** Train the memory encoder, story generator, and sentence generator modules. The sentence and story parsers are running to provide input, but are not being trained.

5. **Epoch 60,000 – 70,000:** Train all modules in a chain.

The training schedule was determined empirically, and is based on a considerable amount of trial and error. Overall, 30 "healthy" DISCERN systems were trained. All of them learned to reproduce the story corpus almost perfectly. On average, 95.6% (SD 0.8%) of sentences and 99.3% of words were reproduced correctly.

The majority of errors that occurred were consistent with errors commonly seen in healthy humans. Seven of the 30 systems jumped once from one to another story that was closely related. All seven jumps occurred in personal stories and stayed within context. All systems sometimes confused story characters that were closely related, most frequently `Stacy` and `Mary`. These confusions showed a strong tendency to stay within context: 86% of the time, a character from the same context was inserted. Lexical errors were rare, and almost always concerned words that were generally used to denote generic conversation topics (as in `I talked to Mary about *books(movies)`). Lexical errors stayed within context 90% of the time. Ungrammatical constructions appeared in less than 0.1% of sentences.

4.5 Conclusion

This chapter described the steps that were completed in order to develop a set of undamaged DISCERN systems as a basis for further experiments. An extensive corpus of input stories was developed and discussed in detail. Based on the words occuring in the stories, word representations were trained to reflect word meanings using the FGREP algorithm. Training methods were developed and then used to train a set of 30 complete DISCERN systems. The recall errors of the resulting systems were analyzed and found to be consistent with normal human performance. The final "healthy" DISCERN systems formed the starting point for the simulations of schizophrenic language described in the next chapters.

Chapter 5

Experiment I: Matching Human Story Recall Data

The foundation of the work reported in this chapter is an experimental study of story recall in patients with schizophrenia, designed and conducted by Ralph Hoffman and his colleagues at Yale (Hoffman et al. 2010) as part of a joint project with this dissertation research. Participants in the study listened to several short stories, and later attempted to recall them as precisely as possible. Insertions, omissions, and recall errors that occurred were recorded and analyzed, creating a detailed characterization of language disturbance in schizophrenia. The scoring methodology of the human study was designed with computational modeling (and the DISCERN model) in mind. Thus, these data present a unique opportunity to tackle one of the main problems that this dissertation is trying to solve, which is to evaluate and distinguish candidate illness mechanisms in a rigorous and quantitative way. Because DISCERN's output is (a simplified version of) human language, equivalent scoring methods can be easily designed for DISCEN. The different illness mechanisms can then be characterized in DISCERN, and compared to human data. The quality of the match can be used to judge which illness mechanism is more likely to cause the observed impairments in patients.

The remainder of this chapter first reviews the study of human story recall, including the experimental design, methodology, and the resulting data. Building on the human study, the next section then describes the methods used to match illness mechanisms to human data, and the computational experiments that were conducted. The results of these experiments are presented next, and are discussed in detail.

5.1 Human Story Recall in Schizophrenia

The participants in the study were 20 healthy controls and 37 patients with schizophrenia. The patients were all relatively stable outpatients under medication, diagnosed psychiatrically with schizophrenia or schizoaffective disorder. These diagnoses were based on DSM-IV criteria (American Psychiatric Association 2000), established using the Comprehensive Assessment of Symptoms and History (CASH; Andreasen 1987). The Yale University School of Medicine Human Investigation Committee approved the human subjects study. Written, informed consent was obtained from all subjects.

In order to examine language-related manifestations of delusions more closely, the patient group was further divided into those who definitely demonstrated evidence of fixed delusions with a plot-like or narrative scheme, and those who showed questionable or no evidence of such delusions. This distinction was made based on the presence of paranoid, grandiose, or religious delusions, and excluded scores for non-fixed delusions (thought broadcasting, thought control, thought insertion, and somatic delusions). Typical examples of such story-like delusions included God choosing the patient to eliminate racial oppression, and the patient being trailed by Homeland Security agents due to allegations of terrorist activities.

The absence of psychiatric diagnosis in healthy controls was confirmed using the non-patient version of the Structured Clinical Interview for the DSM-IV (First et al. 2002). Antipsychotic drug levels of patients were quantified as chlorpromazine equivalents (Davis

Table 5.1: Comparison of patients and healthy controls in the human story recall study, including all individuals completing the seven-day recall task. Patients and controls are generally well matched, although patients had slightly lower WAIS vocabulary scores.

	Age[1]	Gender (M/F)	Parental education (grades)	WAIS scaled vocabulary score
Healthy controls ($N = 20$)	36.6 (9.0)	11/9	13.7 (4.0)	12.2 (3.0)
Patients with schizophrenia ($N = 37$)	41.5 (9.6)	16/21	15.1 (7.6)	9.9 (4.6)
Significance test (two-tailed)	$t(55) = 1.51$, $p = 0.14$	$\chi^2 = 0.72$ $p = 0.40$	$t(55) = 0.77$, $p = 0.44$	$t(55) = 2.04$, $p = 0.046$

[1]mean (stdev)

1974; Woods 2003; Centorrino F 2002). Verbal abilities of all participants were estimated using the Wechsler Adult Intelligence Scale-III vocabulary test (WAIS Wechsler 1987). A comparison of patient and control groups in terms of age, gender, parental education and vocabulary test performance is provided in Table 5.1. Patients and healthy controls were generally well-matched, although patients had somewhat lower vocabulary scores ($p = 0.046$).

In the study, participants were asked to recall three short, pre-recorded stories. The first story was selected from the Chicken Soup for the Soul book series (Cerf 1993), and was chosen because it tended to elicit a sad emotional reaction in a small pilot study with healthy controls. The second story was the "Anna Thompson" story borrowed from the WAIS Logical Memory test, and the third was custom-written to resemble the others. All three stories are reproduced below.

“Flower” Story

In one seat of the bus a wispy old man sat holding a bunch of fresh flowers. Across the aisle was a young girl whose eyes came back again and again to the man’s flowers. The time came for the man to get off. He thrust the flowers into the girls lap. “I can see you love flowers,” he explained, “and I think my wife would like you to have them. Ill tell her I gave them to you.” The girl accepted the flowers and watched the man get off the bus and walk through the gate of an old cemetery.

“Anna Thompson” story

Anna Thompson of South Boston, employed as a cook in a school cafeteria, reported at the police station that she had been held up on State Street the night before and robbed of fifty-six dollars. She had four small children, the rent was due, and they had not eaten for two days. The police, touched by the womans story, took up a collection for her. Anna baked them a cake the following week. Her oldest son, from then on, wanted to be a policeman. Anna never walked down State Street again.

“Hitchhiker” story

I hitched into town. A wispy old man driving a pick-up truck with his frail wife gave me a ride. I sat in the back and watched the tires kick up dust. We stopped and waited for a traffic light. I turned around and peered into the rear window. I hadn’t eaten all day and my eyes came back again and again to a bag of Fritos on the dashboard. The man got out of the truck and walked around to the back. “My wife noticed that you kept looking at the Fritos,” he explained, “and she wanted you to have them.”

Similar to the DISCERN story corpus described previously in Section 4.2, the stories were selected (or, in the case of the Hitchhiker story, written) to have overlapping sctructure and content in order to create the possibility of cross-over or transfer of content between stories. The Flower story and the Hitchhiker story, for example, have similar narrative structure; they also both involve travel and share a story character (the "wispy old man"). All three stories share the theme of a gift being given to a story character.

All participants listened to all three stories in random order. Stories were presented binaurally on headphones. All participants were asked to recall the stories three times: once immediately, once 45 minutes after exposure, and once after seven days. If a subject was unable to recall any element of a particular story spontaneously, he or she was prompted by a the title of the story. Seven-day recall was unannounced in order to prevent rehearsal of stories in the intervening period. The recalled stories of all participants were tape-recorded, transcribed, and then analyzed: errors and insertions were categorized and counted, and overall recall success was scored. A number of outcome variables were calculated based on these results. Among them were ungrammatical constructions, failures of pronoun reference, within-story accretions (misplaced content within a story), and between-story migrations (content from one story intruding into another). The following four variables, however, turned out to be the most reliable and descriptive, with parallel errors of the same type generated by DISCERN under illness conditions, and were therefore used as a basis of comparison with the computational model:

1. **Recall success**

   In order to measure human recall, the three target stories were broken down into sets of kernel propositions. Consider the first sentence of the Flower story: "In one seat of the bus a wispy old man sat holding a bunch of fresh flowers." This sentence was translated into three kernels: (i) a man rode on a bus, (ii) the man was old/aged and frail/weak, (iii) the man was holding/possessed flowers. The Flower story contained 12 kernel propositions, the Anna Thompson story contained 14, and the Gift story contained 10. Successful paraphrases for each kernel was scored if the gist was captured in the rater's judgment. The final recall score was the total number of kernels paraphrased successfully (scored as 1) or partially (scored as 0.5), divided by the number of kernels in all three target stories (36). The recall success variable thus represented the fraction of content correctly reproduced.

2. **Derailments**

   A derailment was scored when a clause (independent or dependent) was produced whose meaning was extraneous to or inconsistent with the target story. An example from the Hitchhiker story, recalled by a patient with schizophrenia: "He got in the truck *and then they stopped for gas*." Stopping for gas was not part or the Hitchhiker story, and the clause was therefore scored as a derailment. The outcome variable for derailments was expressed as a *penetrance score*, meaning the number of derailments divided by the number of correctly recalled propositions. In this way, the variable measures the difficulty in following a consistent story line, expressed as a score that is independent of the number and length of the recalled stories.

3. **Agency shifts**

   These errors comprised a special category of word substitution error that specifically involved story characters. The following example was produced by a patient in the study:

   > "She gave the old man the flowers as a gift."

   In the original story, the old man gave the girl flowers. This segment therefore contained two agency shifts, or agent-slotting errors: one for "She" and one for "the old man." Pronouns referring to people were scored as agency shifts if the implied noun reference was incorrect. In patients, this variable turned out to be linked to fixed delusions. Again, the outcome was expressed as a penetrance score.

4. **Lexical misfires**

   These errors were scored if a word (or word phrase) was replaced by another that filled a similar sentence role, but where meaning was significantly changed from the target text. Examples from human performance include:

   > "Her son was ecstatic" (the son felt good but not ecstatic).
   >
   > "The old man got out of the *car (truck)."

   Agency shifts as defined above were excluded from this category. The outcome variable was lexical misfire penetrance, i.e. number of occurrences divided by recall score.

For each subject, scoring was done by a rater who was not involved in data collection, and who was blind to group, presence or absence of fixed delusions, and subject identification. The interrater reliability obtained for all relevant outcome variables was acceptable (>0.6, see Table 5.1). Pooling data across both groups of human subjects, there was no significant correlation between any of the performance variables, and age, education level, or WAIS-scaled vocabulary, assuming an uncorrected cut-off of $\alpha = 0.05$. Within just the patient group, number of hospitalizations and antipsychotic dose (scored as chlorpromazine equivalents) were also not significantly correlated with any of the performance variables using the same cut-off.

The primary findings of the human study were as follows (Hoffman et al. 2010). First, patients with schizophrenia were significantly less successful ($p = 0.00001$) at re-

Table 5.2: Interrater reliability of all outcome variables is in the acceptable range (> 0.6). Numbers are averaged across immediate, 45-minute, and seven-day story recall data for ten subjects.

Variable	Mean alpha (range)
Propositions recalled	0.98 (0.98 - 0.99)
Agency shifts	0.75 (0.66 - 0.89)
Lexical misfires	0.63 (0.38 - 0.94)
Derailments	0.87 (0. 78 - 0.93)

Table 5.3: Comparative data for healthy controls and patients with schizophrenia-spectrum disorder. Patients showed significant impairment in recall performance, and produced significantly more derailments and agent-slotting errors. Lexical errors did not differentiate the two groups.

	Recall success[1]	Derailment penetrance[1]	Agent-slotting penetrance[1]	Lexical misfire penetrance[1]
Healthy controls	0.67 (0.12)	0.022 (0.072)	0.012 (0.016)	0.026 (0.032)
Patients with schizophrenia	0.41 (0.23)	0.153 (0.178)	0.043 (0.061)	0.033 (.043)
Significance test[4]	$t(55) = 4.9$, $p = 0.00001$	$t(52.1) = 3.9$, $p = 0.0003^{2,3}$	$t(44.8) = 3.0$, $p = 0.004^{3}$	$t(49.3) = 0.7$, $p = 0.49^{3}$

[1] mean (stdev); [2] after sqrt transformation to normalize data;
[3] equal variance not assumed; [4] two-tailed, uncorrected.

producing story content than healthy controls. On average, the recall score was 41% after seven days, or 61% of the average recall score of healthy controls.

Second, derailment behavior was relatively frequent in the patient group, like in the following example produced by a patient in the study: "He got flowers. He looks over at the girl *who has blue eyes*." Penetrance of such insertions differentiated patients from controls robustly ($p = 0.0003$, after square root transform to normalize data).

Third, patients were more likely than healthy controls to produce agency shifts ($p = 0.004$). More interestingly, the group of patients with fixed, story-like delusions made significantly more agent-slotting errors than both healthy controls and patients without these delusions ($p = 0.015$, corrected post-hoc comparison, $\alpha = 0.05$), suggesting that agency shifts may provide a promising model for delusion formation.

One surprising negative finding was that neither lexical misfires nor ungrammatical language differentiated patients and controls. The frequency of lexical misfires was equally low in both patients and healthy controls ($p = 0.49$), and ungrammatical constructions were virtually absent in all study participants.

The human subject data collected in this study present a unique opportunity to evaluate and distinguish the candidate illness models introduced in Section 3.2. The reminder of this chapter describes the methods and the results of a computational study that was designed to do so in a rigorous and quantitative way.

5.2 Methods

In order to evaluate DISCERN and the alternative illness models, a principled way to compare the language produced by DISCERN to that of humans was needed.

While quantitative measures of verbal memory performance have been developed for word list stimuli (Tremont et al. 2000), to our knowledge similar measures do not exist for recall of narratives. Moreover, in the context of altered story recall in schizophrenia, a meaningful comparison should not just measure performance, but instead attempt to capture and compare important aspects of the specific alterations that were found to be relevant.

Based on the findings of the human study summarized in the previous section, such a measure was developed for this experiment. It specifically focuses on the four outcome variables introduced in the previous section: recall success, derailment penetrance, agency shift penetrance, and lexical misfire penetrance.

The primary reason for choosing this set of variables as a basis for comparing human language to DISCERN was that they all indicated some interesting or distictive language behavior in the human study. For example, recall that patients were more impaired than controls in recall success, derailment penetrance, and agent-slotting error penetrance. The lexical misfire penetrance did not differentiate patients from healthy controls, but was used anyway because (1) this was a surprising finding, and (2) different lesions in DISCERN tended to produce different rates of word selection errors (see e.g. Figure 6.3 in the next chapter).

A fifth variable, measuring ungrammatical constructions, was initially used, but was discarded from the analysis because both human study subjects and DISCERN produced virtually no ungrammatical language. The variable was therefore not useful in comparing alternative simulations to the human data.

Other variables were discarded because no equivalent scoring method was available for DISCERN. For example, pronoun reference failures could not be used because DISCERN used only very few pronouns (at least in the current story corpus). Between-story migrations (content from one story intruding into another) could also not be used, because all stories in DISCERN's memory were known, so most errors fell into this category. In humans, on the other hand, intrusions could presumably come from an extremely large number of narrative memories in addition to the target stories used in the experiment. In contrast, the four variables used for comparison could be scored in DISCERN in a straightforward way that was very close or identical to the scoring for human data:

1. **Recall success**

   Sentences in DISCERN were short and contained no dependent clauses, so each sentence was counted as one kernel proposition. Recall success was calculated as the total number of sentences accurately reproduced, divided by the total number of sentences in the corpus of target stories. Stories in DISCERN were much more numerous and often more extended, yielding a total number of 549 kernels.

2. **Derailments**

   In DISCERN, derailments were scored when recall switched from one story to another during recall. The number of derailed sentences was counted and divided by the number of correctly recalled sentences. The resulting penetrance score measured DISCERN's difficulty in maintaining a consistent story line.

3. **Agency shifts**

   Identical to the study of human story recall, agency shifts were scored when DISCERN substituted on story character for another. An example from a DISCERN simulation:

   ```
   I talked to *Stacy(Kate) about *Kate(Stacy).
   ```

   where `Kate` is the mother of the first-person character and `Stacy` is his girlfriend. Like in the example of human agent-slotting errors shown earlier, two story characters are switched, and two agency shifts are scored. The human story recall data suggests that such agent-slotting errors can indicate narrative-type fixed delusions.

4. **Lexical misfires**

   Like in the human subject study, these errors were scored when DISCERN substituted one word for another word that belonged to the same lexical category, unless the substitution was already scored as an agency shift. Examples from DISCERN's output include

   ```
   Vince drove *recklessly(carefully).
   Vito was from *LA(Chicago).
   Kate liked *guns(books).
   ```

   The following word substitutions, on the other hand, were not scored as lexical misfires because they crossed lexical categories:

   ```
   Vince killed *the(Tony).
   I was *wedding(late).
   ```

   Such ungrammatical constructions were rare in this experiment: Most word substitutions in DISCERN involved semantically related words.

Based on these four variables, a measure was developed for how well an illness model matched the human data. It calculated the *goodness of fit* (GOF) for a specific lesioned DISCERN system using a mean square deviation metric (Schobel et al. 2009):

$$\mathrm{GOF}^{C|P}(D, L, s, f) = \sum_{i=1}^{4} \left(\frac{\bar{V}_i^{C|P} - V_i(D, L, s, f)}{\mathrm{SE}(V_i^{C|P})} \right)^2 \tag{5.1}$$

where $\mathrm{GOF}^{C|P}(D, m, f)$ is the goodness-of-fit of a given DISCERN exemplar D with illness model (lesion) L applied at lesion stregth s, and output filter threshold f. GOF is calculated relative to either the group of human healthy controls C or human patients with schizophrenia P. $\bar{V}_i^{C|P}$ is the mean value of the story-recall variable i (either recall success, derailments, agency shifts, or lexical misfires) calculated for the subject group C or P. $V_i(D, L, s, f)$ is the score for that same variable for the DISCERN exemplar D with lesion L at stregth s and filter parameter f. $\mathrm{SE}(V_i^{C|P})$ is the standard error of the mean for variable i. Note that a lower GOF score implies a better fit.

The goodness-of-fit measure was then used to compare alternative illness models to human data. The 30 "healthy" DISCERN models developed in Chapter 4 formed the starting point for this experiment. The initial unlesioned models recalled the entire story corpus reliably, with average sentence-level recall at 95.6% (SD 0.8%).

Each of the eight illness models introduced in Section 3.2 was then tested to determine which one could create the best fit to human data. This comparison was done by the following method, once for the patient group P and once for the healthy control group C.

First, for each illness model L and each of the 30 initial DISCERN systems D, the best-fit combination of lesion strength s and output filter setting f was determined. $\mathrm{GOF}^{C|P}(D, L, s, f)$ was calculated for each combination (s, f) on a 100×1000 grid of lesion strengths and filter settings.

The 100 lesion strengths were spaced equally between zero and a strength where recall performance was close to 30%. Filter settings were varied between zero and a setting that filtered over 90% of the output of unlesioned systems. Preliminary experiments were undertaken to ensure that increasing the resolution of the parameter search further would not alter the outcome significantly.

The result of this parameter search was the best GOF for system D and lesion L, indicating how well lesion L was able to match the human subject group using DISCERN system D. The best-fit GOF data for all 30 systems and all eight lesions were then analyzed to determine which illness model fit the human data best. For the comparison, mixed models were used with best-fit GOF as the response variable, DISCERN systems as the clustering factor, and type of lesion (eight levels) and group (healthy control versus patient) as within-subject factors.

The two illness models that matched patient data best, i.e. working memory disconnection and hyperlearning, were then studied in more detail. The parameter space of both models was expanded by applying the same lesion in a second location. This expansion meant that these models now had three parameters: the lesion strengths s_1 and s_2 for lesions in two separate locations and, as before, the filter setting f.

Hyperlearning was expanded by applying it to the generator modules of DISCERN in addition to the memory encoder module. Network training was adjusted such that s_1 was the learning rate for the memory encoder network, and s_2 was now the learning rate for story and sentence generator modules.

Working memory disconnection was expanded by pruning the story generator's output connections in addition to the connections between context and hidden layer. The lesion parameters s_1 and s_2 were used as pruning thresholds for, respectively, the context-to-hidden layer connections and the hidden-to-output layer connections.

The two expanded illness models were then compared to each other in fundamentally the same way as the two-parameter models. The available computing resources made it possible to expand the parameter space to three dimensions, but the resolution for the lesion strength variables had to be reduced to 40, so that the parameter space was a $40 \times 40 \times 1000$

grid. The statistical approach was the same as previously, using mixed models with the second lesion parameter as an additional fixed effect.

In order to model fixed delusions, the systematicity of DISCERN's agent-slotting errors needed to be investigated, i.e. whether or not the same confusion of a personal-story and a gangster-story character tended to recur in the output stories. This systematicity was measured by generating cross-context errors randomly (using the same base rate for each agent as in the story corpus), and counting how many of the errors were repeats of earlier ones (in the same or opposite direction). The same total number of errors as in the 30 DISCERN exemplars together was generated 10,000 times, and the rate of systematic errors compared to that of DISCERN simulations.

5.3 Results

In the first set of experiments, the eight two-parameter models were compared first to the healthy control group and then to the patient group. A mixed model revealed a significant subject-group $\times$ lesion interaction ($F(7, 203) = 36.7$, $p < 0.0001$).

No illness model had a significant advantage matching the story-recall performance of controls ($F(7, 203) = 1.91$, $p = 0.07$). Figure 5.1A compares the best-fit GOF values obtained by matching each lesion to the healthy control group. In contrast, the eight illness models differed significantly in how well they matched the patients' story-recall performance ($F(7, 203) = 50.5$, $p < 0.0001$). Working memory disconnection and hyperlearning were robustly superior to the other six models in terms of GOF to patient performance (($p < 0.0005$ on paired t-tests) but were not significantly different from each other (Figure 5.1B, Table 5.4).

That the different illness models differ in matching patients but not healthy controls suggests that some models are indeed better able to capture specific aspects of the pathophysiology underlying schizophrenia, rather than the error patterns that humans overall are likely to produce.

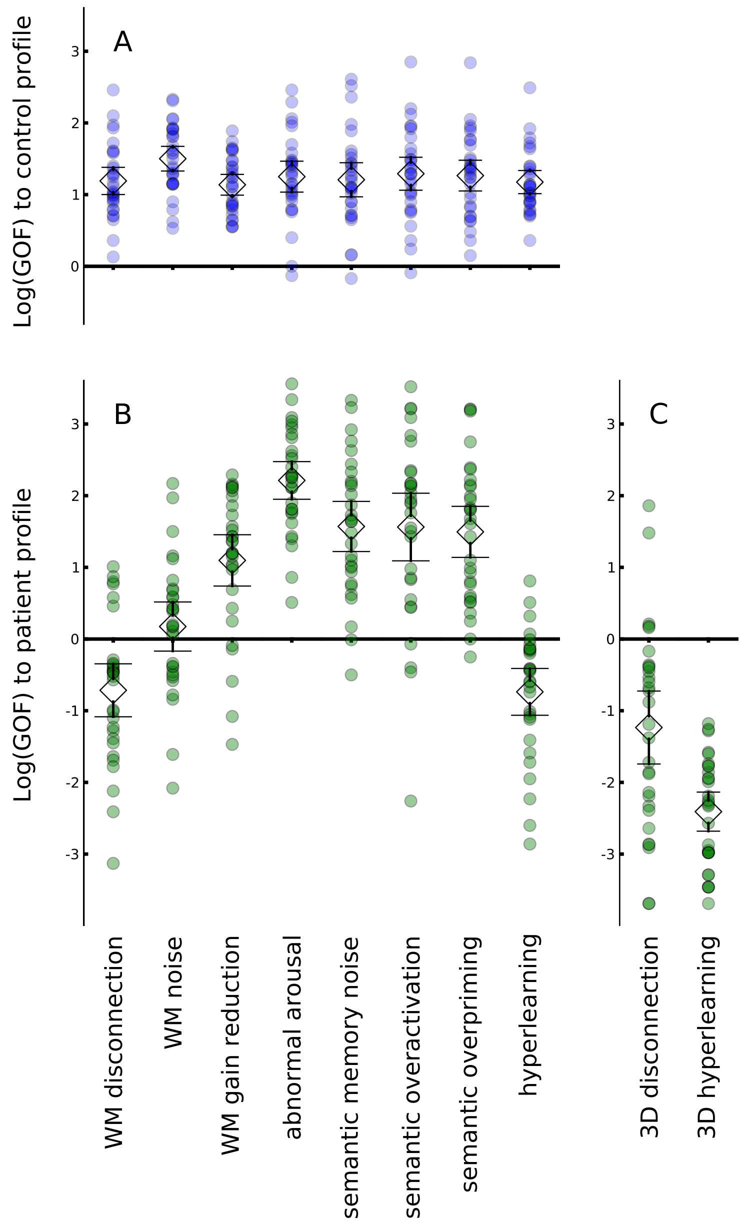


Figure 5.1: Goodness of fit of the 30 DISCERN exemplars to human story-recall data. The plots show each of the eight illness models mapped using a mean square deviation metric to the language profile of controls (A) and patients (B,C). GOF was log converted to normalize distributions; smaller values represent a better fit. All mechanisms were equivalent in matching the controls. However, hyperlearning and WM disconnection fit the patients better than the other mechanisms (Table 5.4). (C) Adding a third parameter to disconnection and hyperlearning models improved GOF to the patient language profile, with hyperlearning fitting significantly better than WM disconnection.

Table 5.4: Pairwise comparisons of optimized GOF for two-parameter hyperlearning and two-parameter disconnection relative to the other six two-parameter models. Comparisons are based on mixed model analysis[1,2].

	Hyperlearning		WM Disconnection	
	t-test	p-value	t-test	p-value
WM noise	3.9	< 0.0001	3.6	< 0.0004
WM gain reduction	7.8	< 0.0001	7.2	< 0.0001
Lowered WM bias	14.7	< 0.0001	13.3	< 0.0001
Semantic noise	9.9	< 0.0001	9.2	< 0.0001
Semantic overactivation	8.2	< 0.0001	7.7	< 0.0001
Semantic blurring	9.5	< 0.0001	8.8	< 0.0001

[1]df = 203, all pairwise comparisons favored 2D hyperlearning and disconnection over other models; numbers in parentheses correspond to mechanism code illustrated in Figure 3.7; [2]comparison of 2D hyperlearning vs 2D WM disconnection in terms of optimized GOF with patient data was nonsignificant (t=0.09).

In the second experiment, these two illness models were further studied by adding a second lesion parameter to each, resulting in a three-parameter model. The third parameter improved the overall GOF to patient data significantly ($F(1, 29) = 37.3$, $p < 0.0001$). A significant lesion $\times$ parameter interaction was detected ($F(1, 29) = 10.3$, $p = 0.003$), with three-parameter hyperlearning producing a significantly better fit to the patient profile than the three-parameter disconnection model ($t(29) = 4.2$, $p = 0.0002$).

In order to evaluate how well the best-fit illness model was able to reproduce the actual language behavior of patients, the content of the language produced by the three-parameter hyperlearning model was analyzed in detail. First, with regard to derailments, midstream jumps from one story to another occurred in a highly systematic fashion in both the model and the patients (who were in relatively stable condition during the study). Most striking was that emotional valence was retained from the pre- to the post-derailment story in 90.1% of the 30 simulations. Moreover, personal stories tended to derail to other personal stories, and gangster stories tended to derail to other gangster stories. Overall, only 15.1% of derailments violated context. A typical example of an autobiographical derailment was initiated when DISCERN recalled story about the self meeting his mother, Kate, for a drink, and recounting some of her attributes:

```
Kate had a ponytail.
Kate drove a nice car.
Kate liked books.
```

DISCERN then switched into another personal story with similar content and structure, in which the self meets his girlfriend Stacy for a drink:

```
I talked to *Kate(Stacy) about *guns(books).
I like *baseball(books).
I liked talking with Stacy.
```

Note that the derailment includes an agent slotting error that is consistent with the original story. In patients with schizophrenia, derailments are often accompanied by such errors. The two lexical misfires both substitute closely related words – `guns`, `book`, and `baseball` all appear in the corpus mainly as topics of conversation, only distinguishable by the people who talk about them. This, also, is consistent with the kind of errors seen in patients. Interestingly, the original story continues with

```
I talked to Kate about Stacy
I liked Stacy
[...]
```

It seems that the imminent appearence of Stacy in the target story may have caused the confusion between the two stories, suggesting that, at least in this case, the discourse was derailed by a confusion between two alternative discourse plans with similar content, rather than more low-level fluctuations in network error. Also of note is that both the "meeting Kate for a drink" and the "meeting Stacy for a drink" stories (#18 and #16 in the story corpus) were assigned the same emotional valence, "+,", so DISCERN could not rely on this information to guide story recall.

Second, in regard to delusions, recall that the human data demonstrated that patients with fixed, story-like delusions are more likely than others to produce agent-slotting errors. This suggests a possible theory of delusion formation: delusional ideas and stories could emerge when agents (specifically the self) cross over into impersonal stories, creating spurious and possibly frightening narratives that are then remembered as lived experience. In order for this agent-slotting model of delusions to work, agency shifts need to cross contexts in a systematic fashion.

The hyperlearning systems produced cross-context agent-slotting errors in a highly systematic fashion: the same two characters, one from an autobiographical story and the other from a crime story, were interchanged on average 2.5 occasions (range 0-10) per exemplar. For example, the following output was generated by one hyperlearning exemplar when recalling Story #2:

```
*Vito(Joe) was in his 30s.
Joe was a doctor.
Joe worked in New York.
Joe was my boss.
```

In the story corpus, Vito is the boss of the Mafia gang in the gangster context, while Joe is the boss of the self in autobiographical stories. Later, for Story #10, the same DISCERN system produced

```
Vito was a famous gangster.
Vito was the boss of Tony.
Tony hated Vito.
Tony feared *Joe(Vito).
```

The same confusion occurred again while recalling Story #27:

```
Vince went to Starbucks.
Vince sat at a table.
Vince liked Vito.
Vince feared *Joe(Vito).
Vince accepted the order.
Vince drove to City Hall.
Vince was scared.
Vince drove carefully.
Vince entered City Hall for murder.
```

In the original version of this story, Vince, a Mafia hitman, received an order from Vito to kill Tony, another gangster. Again, the Joe reference is substituted for Vito. In human terms, it seems plausible that a tendency to systematically confuse Joe and Vito could lead to a delusional belief that one's boss is really a Mafia boss. The frequency with which the same pair of agents was confused across contexts by best-fit three-parameter hyperlearning models was robustly greater than expected by chance ($p < 0.00001$ in the randomization test). This finding demonstrates that these models confuse story characters in a systematic fashion, suggesting that the hyperlearning model may indeed capture a central aspect of the emergence of delusinoal narratives in humans.

5.4 Conclusion

Eight alternative illness models based on DISCERN were compared to human language and memory performance based on a combination of four measures: recall success, derailment errors, agent-slotting errors, and lexical misfire errors. A mean square deviation metric was used to estimate how well the different illness models were able to match both healthy humans and patients with schizophrenia.

The main result of this experiment is that whereas all eight illness mechanisms were equivalent in matching the story-recall profile of healthy controls, hyperlearning was significantly better than the others in matching the story-recall profile of patients. Taken together, these findings suggest that the hyperlearning model captures specific aspects of altered brain processes underlying schizophrenia, rather than nonspecific sources of error-proneness demonstrated by human subjects overall.

The language resulting from hyperlearning resembles that produced by patients also qualitatively. Derailments are highly non-random and appear to be caused by distortions of discourse planning rather than lower-level network damage, similar to derailment behavior seen in patients. Delusions are modeled by systematic agent-slotting errors that cross over between contexts, which accounts for human data and suggests that delusions could form when characters (specifically the self) cross over into impersonal stories, creating spurious narratives that are then remembered as lived experience.

In summary, these findings demonstrate that using the DISCERN model, it is possible to evaluate and compare alternative illness models to human performance in a rigorous and quantitative way. Hyperlearning emerged as the best illness model, and was able to capture the language impairment observed in patients both quantitatively and qualitatively.

However, recall that the patients who participated in the study were in relatively stable condition, so the best-fit simulations produced more or less intact stories rather than the more severely impaired language often seen in acute psychosis. An attempt to simulate these more severe language abnormalities was the logical next step.

Chapter 6

Experiment II: Psychotic Language

Psychosis is the hallmark of schizophrenia (Kapur 2003), and the positive symptoms of schizophrenia are of special interest to this work. First, they are among the most distinctive, and therefore among the easiest to identify in the language of patients (or DISCERN). They are central to the definition and diagnosis of schizophrenia, and they tend to dominate in the early stages of the disorder, where models of intervention could be most useful. Consequently, this chapter aims to recreate the language disturbance characteristic of acute psychosis, and evaluates the ability of the illness models to do so.

6.1 Characterizing Schizophrenic Language

There is not much quantitative data available about the precise language disturbance profile of actively psychotic patients with schizophrenia. However, diagnostic criteria (American Psychiatric Association 2000; Andreasen 1984, 1987) and the literature that exists (e.g. Andreasen 1979; Hoffman et al. 1986; Appelbaum et al. 1999; Covington et al. 2005; Kuperberg 2010) provide a relatively clear qualitative picture of the kind of language disturbance that would produce a compelling model of psychosis in schizophrenia.

Since delusions and disorganized speech are the major psychotic symptoms of schizophrenia that are expressed directly via language, the experiments in this chapter will focus on these two symptoms. Disorganized language is often characterized by derailments, where patients tend to jump from one story to the next, creating a confusing, fragmented narrative composed of vaguely related "discourse shards" (Hoffman et al. 1986). Andreasen (1979) also reported several other signs of language disorganization that are at least "moderately common" in schizophrenia, including poverty of speech, tangentiality, and perseveration.

Delusions in schizophrenia tend to stay fixed over time, and often have paranoid or grandiose content (Harrow et al. 2004). A majority of these delusions insert the patient or persons known to the patient into a rigid, implausible or bizarre narrative schema (Appelbaum et al. 1999; Vinogradov et al. 1992). The human subject study described earlier confirmed that such agency shifts can be a sign of story-like delusions in patients, and the computational study in the previous chapter suggested that agency shifts may be a mechanism by which delusional narratives emerge.

At the sentence or word level, schizophrenic language tends to be mostly intact: Word pronounciation, morphology, and syntax are all normal or nearly so (Covington et al. 2005; Hoffman and Sledge 1988), even though some syntactic impairments have been reported (Hoffman and Sledge 1988).

Another reported characteristic of schizophrenia is insensitivity to context. In terms of language abnormality, this means that words may often appear incongruous or inappropriate given the surrounding language. Such "context impairments" have also been proposed as the underlying reason for language disorganization in general (Cohen and Servan-Schreiber 1992; Kuperberg 2010). One way to distinguish alternative hypotheses, then, would be to look for errors that could have been easily avoided by using contextual cues, but not otherwise. The two separate contexts (personal and gangster stories) in the story corpus (Section 4.2) were designed to provide opportunities to observe such errors. For example the self's boss `Joe` and the gangster boss `Vito` are similar in many ways, but are easily distinguished by the story context in which they appear.

Taken together, these features of schizophrenic language suggest that the processing of overall discourse structure is more impaired than sentence-level language processing: Derailments and delusions can be seen as failures of global story structure and content, while locally, structure and meaning remain relatively intact. A successful model should be able to capture these characteristics, i.e. errors should be failures of context and discourse rather than break-down of syntax and lexical access. The experiment described in this chapter evaluates and compares the ability of different illness models to recreate these features of schizophrenic language.

6.2 Methods

The experiments in this chapter were conducted using ten heathy DISCERN systems that were all trained in the same way but using different random initial connection weights. As before, the story corpus, lexicon, and training schedule developed in Chapter 4 were used. The average sentence-level recall rate of the intial undamaged systems was 96%.

The ten DISCERN systems were subjected to varying degrees of lesion damage using the same eight simulated illness mechanisms investigated in the previous experiment. Additionally, hyperlearning was applied to DISCERN's generator modules, for a total of nine lesions. The strength of each lesion was increased gradually, starting at zero and adding small increments until the resulting recall performance was approximately 30%. Since the present experiment focuses on psychotic language behavior rather than negative symptoms, DISCERN's output filter was not used. Table 6.1 provides an updated summary of all lesions. For brevity, abbreviations from this table are used to refer to each lesion in this chapter (e.g. "HLM" for "hyperlearning applied to the memory encoder module").

Errors were counted and classified automatically, using the same error categories as in the previous experiment, i.e. ungrammatical sentences, derailed language, lexical errors,

Table 6.1: Candidate illness models used in experiment II.

Abbreviation	Simulated illness mechanism	Lesion parameter
SB	Semantic Blurring	Radius
WMD	Working Memory Disconnection	Pruning threshold
SN	Semantic Noise	Noise variance
WMG	Working Memory Gain Reduction	Gain change
EA	Excessive Arousal States	Added bias
SO	Semantic Overactivation	Added bias
WMN	Working Memory Noise	Noise variance
HLM	Hyperlearning (Memory Encoder)	Learning rate
HLG	Hyperlearning (Generator Modules)	Learning rate

and agency shifts. In addition to the basic categories of errors, a more detailed analysis of the patterns and structure of agency shifts produced by different lesions was conducted for this experiment. First, agency shifts were further divided into those that crossed contexts (i.e. a gangster being inserted into a personal story or vice versa) and those that did not. Second, a separate subcategory of agency shifts was created for *self insertions*. If a story character was replaced by the self character, this was counted as a regular agency shift, but also separately as a self insertion.

Third, agency shifts were divided into unique vs. repeated agency shifts. To be counted as a repeat, the same substitution (i.e. the same combination of intruding and replaced character) had to have occurred previously within the same story. How many of the observed agency shifts were repeated was used as a measure for the stability of patterns of agency shifts.

Fourth, a weighted entropy metric was developed, intended as an alternative measure of the consistency and predictability of agency shifts. Intuitively, the entropy measure computes how hard it is on average to predict agents in the output text, given previous knowledge of the "correct" agents in the story corpus:

$$\sum_{i \in A} -\frac{n_i}{|A|} \sum_{j \in A} P(i,j) log_2(P(i,j)),$$

where A is the set of all agents in the story corpus, n_i is the number of occurrences of agent i in the story corpus, and $P(i, j)$ is the relative frequency with which agent j (as opposed to other characters) is substituted for agent i. The resulting measure of entropy is expressed in bits. For example, if no agency shifts occur, the weighted entropy would be zero, and if exactly half of the instances of every agent were replaced by one particular agent, the weighted entropy would be 1 bit.

6.3 Overview of Results

All illness models reduced the recall performance and led to significant distortion of DISCERN's output language. This section analyzes the differences in the types and patterns of recall errors across the alternative lesions. The goal is to determine which lesions are able to produce language distortions that are both plausible models of psychotic symptoms and consistent with the language observed in patients with schizophrenia.

Figure 6.1 gives an overview of the type of language distortion that was observed following the different lesions. For each lesion, recall errors over a range of increasing damage are broken down into the four error categories. Error percentages are averaged over the ten individual DISCERN systems.

A few interesting differences are immediately obvious from these plots. First, the two kinds of hyperlearning lesions seem to produce patterns of recall errors that are strikingly different from each other. When hyperlearning is applied to the generator modules (HLG), agency shifts dominate other errors; when it is applied to the memory encoder module (HLM), most recall errors are derailments. No other lesions produces output that is dominated by one kind of error in this way.

Second, both forms of hyperlearning produce virtually no ungrammatical language. In contrast, all other lesion produce a substantial percentage of ungrammatical propositions. To a lesser degree, the same is true for lexical insertion errors.

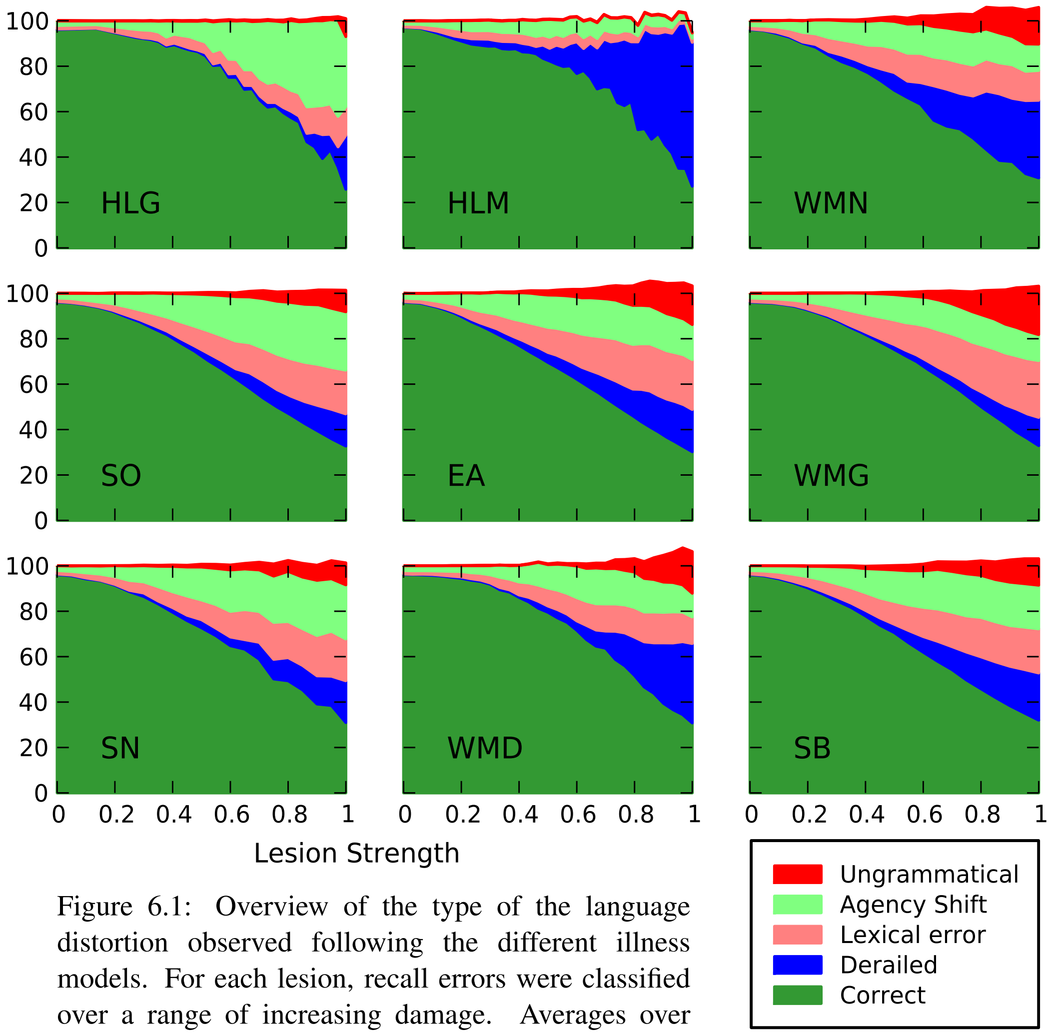


Figure 6.1: Overview of the type of the language distortion observed following the different illness models. For each lesion, recall errors were classified over a range of increasing damage. Averages over ten individual DISCERN systems are shown. The error patterns produced by the hyperlearning lesions (HLM and HLG) are qualitatively different from each other and from the other lesions: HLM produces mostly derailed language, and HLG produces mostly agency shifts. In contrast, error patterns produced by the other lesions are relatively uniform, are not dominated by one kind of error, and contain many grammatical errors.

Finally, the patterns of all lesions that do not involve hyperlearning (called “non-HL lesions”) are relatively uniform, although some produce more ungrammatical language or derailments than others. Differences may still emerge from a more detailed analysis, but it appears that the most promising qualitative differences exist between HLG, HLM, and the non-HL lesions.

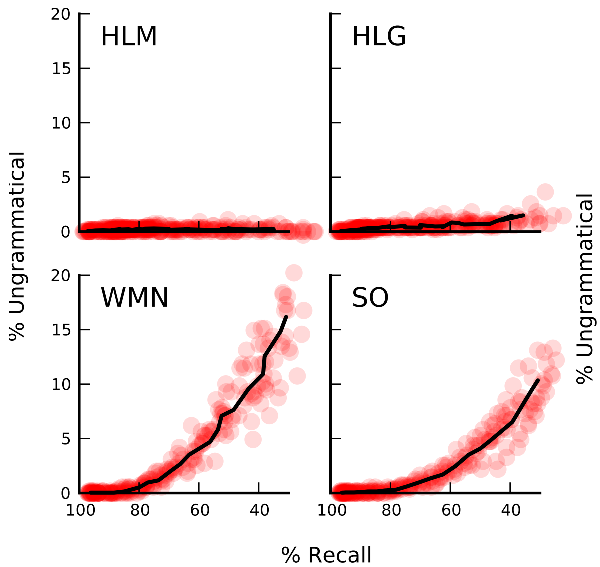
Note that DISCERN’s recall performance degrades differently and generally non-linearly dependng on the lesion. It would therefore be difficult to find “eqivalent” lesion stregths at which to compare the alternative lesions fairly. In order to avoid this problem, reduced recall performance is used as a measure of impairment, and comparisons are generally made at equal levels of recall. Where it is necessary to pick a single level of impairment, comparisons are made at 40% recall, which is typical for patients with schizophrenia (see the human study described in Chapter 5). When measures of impairment other than recall are used, they are motivated separately in each case.

The remainder of this chapter takes a closer look at the differences apparent in Figure 6.1, and examines the impact of different error patterns on the actual language produced by DISCERN. The main goal is to determine if any lesions are able to produce language abnormalities suggestive of (and consistent with) the psychotic symptoms of schizophrenia.

All data used in this experiment, including the output stories generated by all illness models, can be found at http://nn.cs.utexas.edu/?schizo.

6.4 Grammatical and Lexical Errors

Patients with schizophrenia generally produce language with relatively intact syntax and morphology, although exceptions have been reported. High levels of grammatical errors would therefore be problematic in the output language generated after a lesion. Figure 6.2 shows the differences in ungrammatical language that were already visible in Figure 6.1 more clearly. On the left, the effects of four representative example lesions on the level of grammatical errors are shown as recall performance drops. Clockwise, starting on



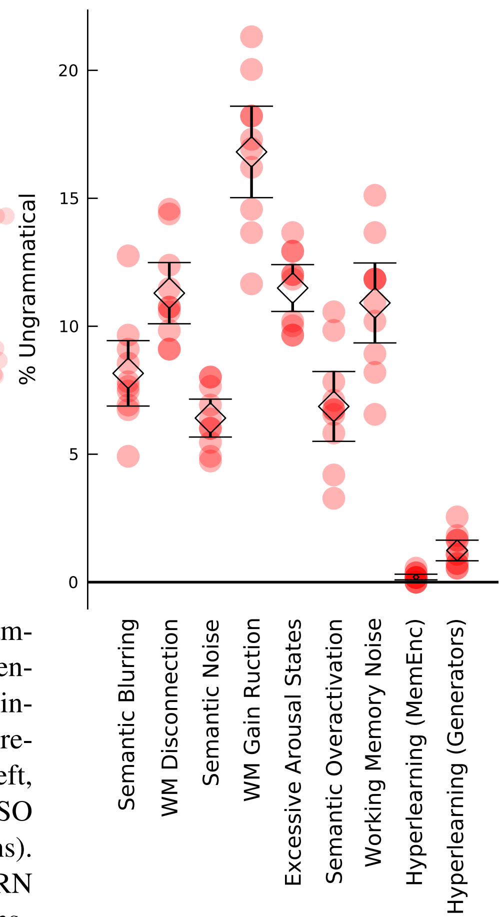


Figure 6.2: *Left:* The percentage of ungrammatical language produced by four representative lesions is shown as lesion damage increases and recall performance drops in response. Clockwise, starting on the top left, these lesions are HLM, HLG, WMN, and SO (see Table 6.1 on page 101 for abbreviations). Solid lines are averages over ten DISCERN systems. Dots are individual lesioned systems.
Right: The percentage of ungrammatical language for all lesions. For a fair comparison, the intensity of each lesion was adjusted so that recall performance was as close as possible to 40% in each case. All lesions except HLM and HLG produce high levels of ungrammatical language, which is not usually seen in schizophrenia.

the top left, these lesions are HLM, HLG, working memory noise (WMN), and semantic overactivation (SO). For the larger plot on the right, the intensity of each lesion was adjusted for each DISCERN system such that recall performance was as close as possible to 40%.

The levels of ungrammatical language observed in non-HL lesions are inconsistent with the language observed in patients with schizophrenia, suggesting that these lesions are not plausible illness models. However, using the output filter described in Section 3.1.5,

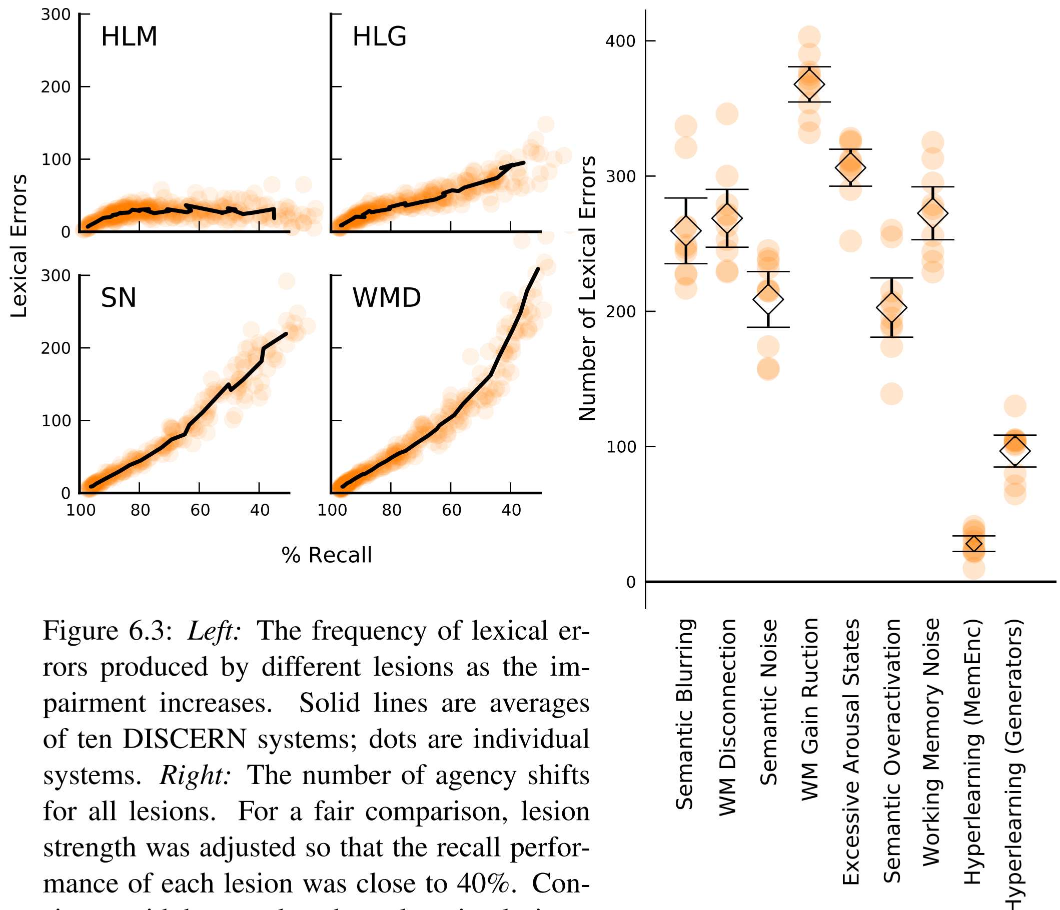


Figure 6.3: *Left:* The frequency of lexical errors produced by different lesions as the impairment increases. Solid lines are averages of ten DISCERN systems; dots are individual systems. *Right:* The number of agency shifts for all lesions. For a fair comparison, lesion strength was adjusted so that the recall performance of each lesion was close to 40%. Consistent with human data, hyperlearning lesions, especially HLM, produce few lexical errors.

ungrammatical language can be effectively eliminated at the cost of reduced language output. This result is nevertheless problematic, because patients with schizophrenia do not universally talk less. Non-HL lesions would therefore have trouble accounting for these patients.

Similarly, Figure 6.3 illustrates that both hyperlearning lesions produce fewer lexical errors, i.e. recall errors where one word is substituted for another. This category excludes errors that exchange story characters (agency shifts), which will be discussed in more detail below.

The following snippet of DISCERN's output language illustrates the impact that grammatical and lexical errors can have on stories recalled by DISCERN. It was produced by a DISCERN system after WM disconnection (threshold 0.2). The overall sentence-level recall of the system was 51%. Word substitutions are marked with an asterisk (*), and are followed by the correct word in parentheses.

```
The Police investigated the bombing at City-Hall.
The Police looked for evidence.
The Police found that Tony bombed City-Hall.
The Police was after *to(Tony).
The Police was after Tony.
The Police thought that *the(Tony) bombed *airport(City-Hall).
The Police wanted to arrest *the(Tony).
The Police found that *to(Tony) *LA(was) *St-Mary's(in).
The Police planned to arrest *I(Tony) in New-York.
[...]
```

In this example, distorted sentence constructions like "The Police found that to LA St-Mary's" make the text appear non-sensical (St. Mary's is the hospital where the self works). In contrast, the following was produced by the same DISCERN system reproducing the same story after hyperlearning was applied to the generator modules (strength=0.3). Even though the overall recall is about the same ($\tilde{5}0\%$), the resulting language is much more coherent and grammatical:

```
The Police investigated the bombing at City-Hall.
The Police looked for evidence.
The Police found that Tony bombed City-Hall.
The Police was after *I(Tony).
The Police was after *I(Tony).
The Police thought that *I(Tony) bombed City-Hall.
The Police wanted to arrest *Joe(Tony).
The Police found that *I(Tony) was in New-York.
The Police planned to arrest Tony in New-York.
[...]
```

These short examples also illustrate a more general difference between hyperlearning and all other lesions: in the first example, syntax and meaning break down in a local way, leaving the overall story random and meaningless. In the second example, the content is changed in a way that changes the overall meaning of the story, but leaves it locally coherent. Errors of this kind, suggesting global errors of context and organization rather than local, sentence-level breakdown are pervasive in schizophrenia. In the following two examples, these differences stand out even more clearly. Again, they are taken from the same DISCERN system recalling the same story (story #12, script 1). The first was produced following WMD (strength=0.27):

```
I went to Moe's-Tavern.
I sat at the counter.
I ordered *coffee(beer).
I drank the *bombed(beer).
I met *airport(Joe) at Moe's-Tavern.
```

The second example was produced using HLG (strength=0.36):

```
*Tony(I) went to Moe's-Tavern.
*Tony(I) sat at the counter.
*Tony(I) ordered beer.
*Tony(I) drank the beer.
*Tony(I) met Joe at Moe's-Tavern.
```

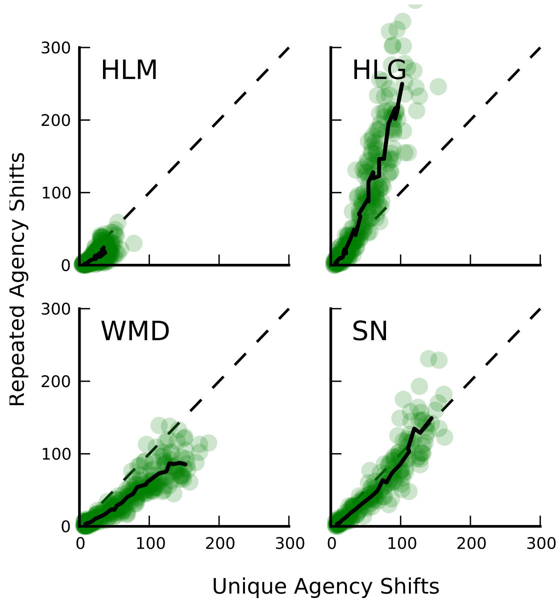
Again, WMD produces ungrammatical and apparently meningless language, while the language generated by HLG remains locally consistent and coherent, while the overall meaning is distorted.

6.5 Agency Shifts and Delusional Language

A plausible model of delusional language needs to demonstrate how delusional ideas and narratives could emerge from normal experience and shared cultural stories. As mentioned earlier, one possible mechanism that has been proposed for delusion formation is that patients with schizophrenia insert themselves or persons in their life into complex imaginary narratives, creating spurious memories that can acquire the same force as lived and remembered reality. The concept of delusion formation based on such *agency-shifts* is supported by fact that patients with fixed, story-like delusions are more likely to confuse agents when recalling stories (Section 5.1).

Of all lesions investigated, only hyperlearning produced a compelling model of delusional language. Applied to the generator networks, hyperlearning robustly produced stable patterns of agency shifts where characters migrated between stories in a highly consistent way. Like in the examples shown in the previous section, grammar and local structure stayed largely intact, while emerging patterns of agency shifts produced meaningful (but distorted) new narratives. For example, in the following output text the gangster boss Vito is replaced by the self's own boss Joe:

```
*Joe(Vito) drove recklessly .
*Joe(Vito) was pulled-over by a cop.
The cop asked *Fred(Vito) for his license.
*Joe(Vito) gave his license to the cop.
The cop checked the license.
The cop arrest(ed) *Joe(Vito) for bombing.
*Joe(Vito) was accused of *murder(bombing).
*Joe(Vito) was brought before the court.
*Joe(Vito) had a good lawyer.
The court cleared *Joe(Vito) of bombing.
*Joe(Vito) walked free.
```



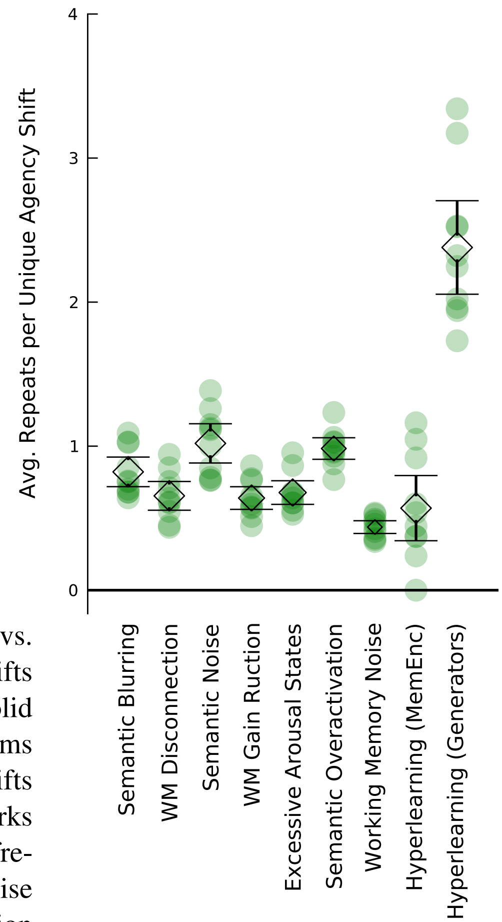
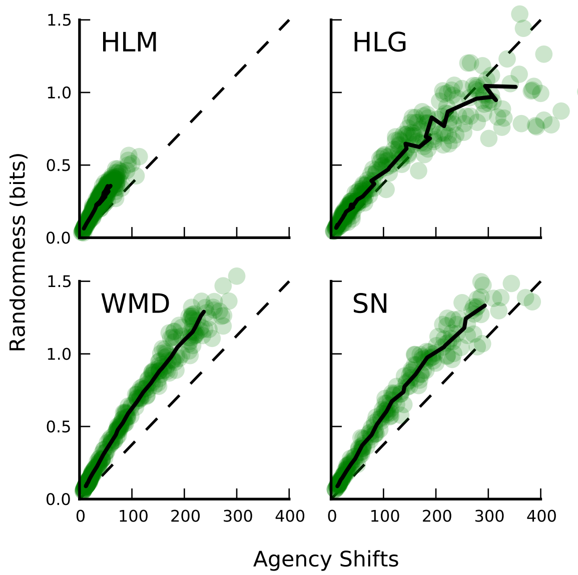


Figure 6.4: *Left:* The number of unique vs. repeated (within the same story) agency shifts following four representative lesions. Solid lines are averages of ten DISCERN systems as the lesion damage increases. Agency shifts following hyperlearning of generator networks (HLG) tend to be repeated much more frequently than those caused by semantic noise (SN) or working memory disconnection (WMD). Hyperlearning applied to the memory encoder (HLM) produces few agency shifts. *Right:* Average number of repeats per unique agency shift for all nine lesions. Again, character substitutions following HLG are repeated much more frequently, causing stable patterns of agency shifts suggestive of delusional content.

In this way, characters from the self's personal context often intrude into gangster stories. Combined with the substitution of "murder" for "bombing" this pattern of errors creates a new story whose content has little to do with any input story DISCERN has ever seen. It is easy to imagine how such stories, if they are remembered as real, could lead to the formation of complex and frightening delusional ideas.



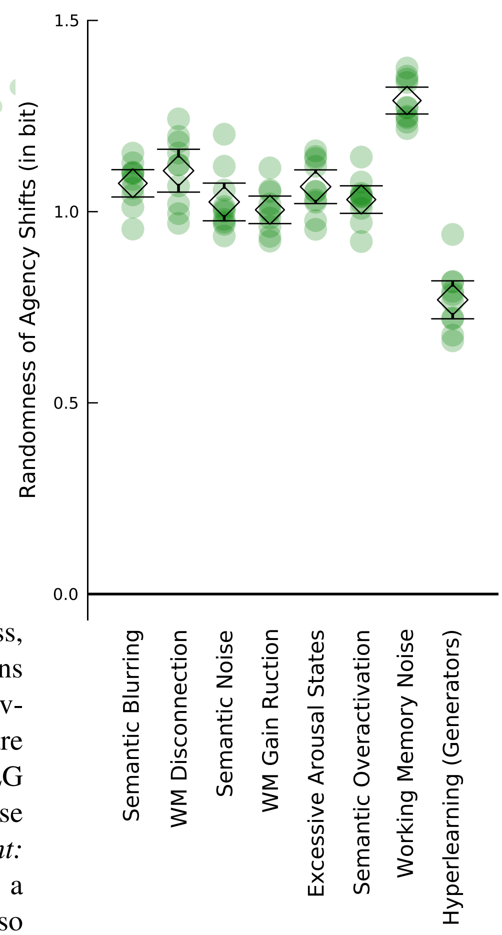


Figure 6.5: *Left:* The entropy, or randomness, of agency shifts produced by different lesions as the impairment increases. Solid lines are averages of ten DISCERN systems; green dots are individual systems. Agency shifts due to HLG have lower entropy than SN or WMD because they follow a more consistent pattern. *Right:* Entropy of agency shifts for all lesions; for a fair comparison lesion strength was adjusted so that each lesion produced close to 200 agency shifts. Again, HLG producesagency shifts with lower entropy, suggesting more consistent and predictable error patterns. HLM was omitted from this plot because it never produces 200 agency shifts, even with very high lesion damage.

Consistent, "delusion-like" patterns of agency shifts such as these are pervasive in the output of HLG, and very rare or absent in other lesions. Figure 6.4 shows that agency shifts following HLG are more consistent than those generated by other lesions: they are repeated several times on average within the same story. Agency shifts following other lesions, including semantic noise (SN) and working memory disconnection (WMD), tend to be repeated much less frequently. Interestingly, HLM produces very few agency shifts, but also does not produce many repetitions.

This finding is confirmed when the entropy of agency shifts is used as an alternative measure of consistency. Figure 6.5 shows that HLG produces agency shifts that have lower entropy than those produced by other lesions, suggesting that they follow a more predictable pattern. To make a fair comparison possible in this case, the lesion strength was set separately for each system and each lesion so that close to 200 agency shifts were produced. Comparisons were not done at equal recall performance in this case because HLG produces

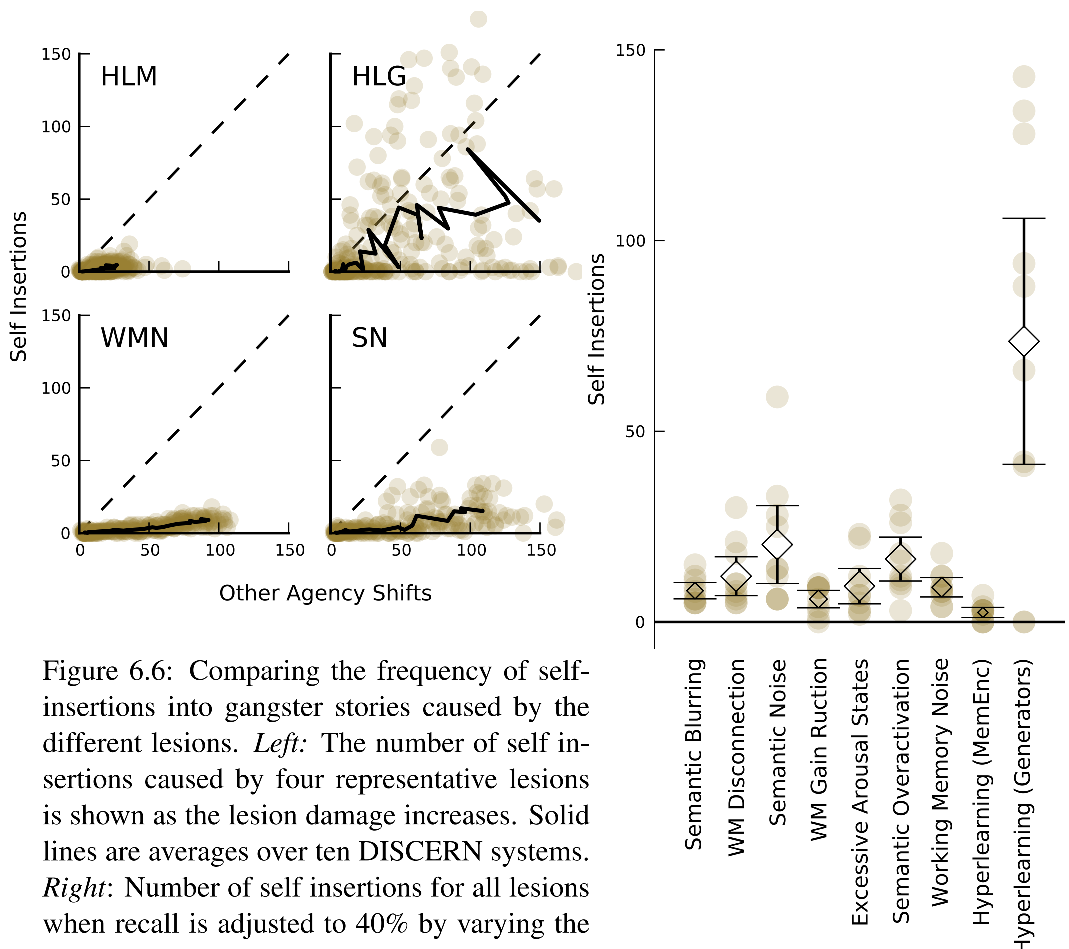


Figure 6.6: Comparing the frequency of self-insertions into gangster stories caused by the different lesions. *Left:* The number of self insertions caused by four representative lesions is shown as the lesion damage increases. Solid lines are averages over ten DISCERN systems. *Right*: Number of self insertions for all lesions when recall is adjusted to 40% by varying the amount of lesion damage. Only HLG shows a pronounced tendency to insert the self into gangster stories. Note that in some systems HLG does not cause any self-insertions, while in others, most (up to 80%) of all agency shifts in gangster stories are self insertions.

more agency shifts than other lesions at equal recall, and the entropy depends in large part on the number of agency shifts.

Patients with story-like delusions often insert themselves into the delusional narrative. For example, a delusional patient is more likely to believe that he himself, rather than someone else, is being followed by the CIA. The agency shifts produced by HLG show the same pattern. Often, when the self takes over for a gangster character the shift remains stable throughout the remainder of the story:

```
The Police investigated the murder at City-Hall.
The Police looked for evidence.
The Police found that Vince killed Bob.
The Police was after Vince.
The Police was after Vince.
The Police thought that Vince *met(killed) Bob.
The Police wanted to arrest Vince.
The Police found that Vince was in New-York.
The Police planned to arrest Vince in New-York.
Vince wanted to go to *Chicago(LA).
*I(Vince) entered his car.
*I(Vince) drove to the airport.
*I(Vince) was *on-time(scared).
*I(Vince) drove recklessly.
*I(Vince) was pulled-over by a cop.
The cop asked *me(Vince) for his license.
*I(Vince) gave his license to the cop.
The cop checked the license.
The cop arrested *me(Vince) for *wedding(murder).
*I(Vince) was accused of murder.
*I(Vince) was brought before the court.
*I(Vince) had a good lawyer.
The court convicted *me(Vince) of *wedding(murder).
*I(Vince) went to jail.
```

In this version of story #28, the self replaces the gangster Vince. Interestingly, the word "murder" is replaced with "wedding", a word that is otherwise exclusively used in personal stories. This lexical error is repeated across two separate scripts. Note also that the self is inserted not only through the word "I", but also through "me" where appropriate. That pronouns like "me", "my", and "his" are mostly (not always) correctly adapted is further evidence that HLG distorts and impairs story processing in DISCERN at the level of actual content rather than that of superficial language production.

Another interesting aspect of the patterns of self insertions produced by HLG is that they tend to recur multiple times within the output of the same DISCERN system. Many systems never insert the self character at all; however, those who do tend to do so repeatedly over multiple stories. Figure 6.6 illustrates this tendency: not only are self insertions much more frequent for HLG on average than for any other lesion – the distribution within the HLG lesion is very broad, i.e. some systems produce very few self insertions, while in some systems, up to 80% of agency shifts in gangster stories insert the self. The DISCERN system that produced the example above was solidly in the latter category: It also produced the following version of story #21, where the self takes over for the gangster Tony:

```
[...]
*I(Tony) was a *doctor(gangster) .
Tony worked for the Mafia.
*I(Tony) worked in New-York.
*I(Tony) hated *my(his) job.
*I(Tony) was a bad gangster.
*I(Tony) wanted to go to City-Hall.
*I(Tony) entered his car.
*I(Tony) drove to City-Hall.
*I(Tony) was *on-time(scared).
*I(Tony) drove *recklessly(carefully).
*I(Tony) entered City-Hall for *wedding(bombing).
Tony bombed City-Hall.
```

```
The *meeting(bombing) was a success.
*Vince(Tony) made a phone-call.
*I(Tony) smoked a cigarette.
*I(Tony) went to Moe's-Tavern.
*I(Tony) sat at the counter.
*I(Tony) ordered beer.
*I(Tony) drank the beer.
*I(Tony) met no-one at Moe's-Tavern.
*I(Tony) ordered more beer.
*I(Tony) got very drunk.
*I(Tony) had a bad time.
```

In addition to the self insertions, the confusion between different types of events recurs, this time including wedding, bombing, and meeting. This type of error is especially interesting because wedding, bombing, and murder (not meeting) are highly context-specific. The fact that DISCERN confuses these concepts in a consistent manner suggests that it specifically misreads cues concerning story context. In story #27 (recalled by the same system) the self takes over for Vince once again, and the same confusion between wedding, meeting, and murder re-emerges:

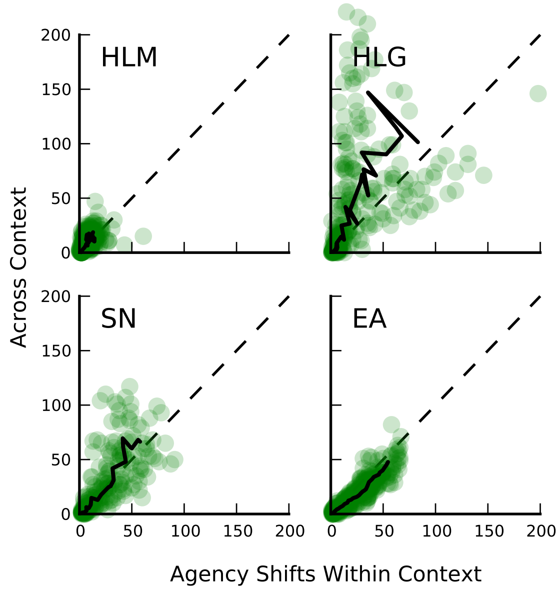
```
[...]
*I(Vince) wanted to go to City-Hall.
*I(Vince) entered his car.
Vince drove to City-Hall.
*I(Vince) was *on-time(scared).
*I(Vince) drove *recklessly(carefully).
*I(Vince) entered City-Hall for murder.
*I(Vince) killed Bob.
The *meeting(murder) was a success.
Vince made a phone-call.
*I(Vince) smoked a cigarette.
```

```
*I(Vince) wanted to go to City-Hall.
*I(Vince) entered his car.
*I(Vince) drove to *to(City-Hall) *Four-Seasons(_).
*I(Vince) was scared.
*I(Vince) drove carefully.
*I(Vince) entered City-Hall for murder.
*I(Vince) killed Bob.
The *wedding(murder) was a success.
*I(Vince) made a phone-call.
*I(Vince) smoked a cigarette.
```

The three examples of self-insertions above were all taken from the output of the same DISCERN system (#4) after the same lesion at the same strength was applied to it (HLG at 0.35). The examples were selected in this way to demostrate that delusion-like error patterns are often shared across stories within the same system. However, error patterns like these are not specific to any DISCERN system or lesion strength. The following output, for example, was produced by a different DISCERN system with a weaker lesion:

```
[...]
*I(Tony) was scared.
*I(Tony) drove carefully.
*I(Tony) entered City-Hall for bombing.
*I(Tony) bombed City-Hall.
The *wedding(bombing) was a success.
*I(Tony) made a phone-call.
*I(Tony) smoked a cigarette.
[...]
```

Repeated patterns such as these are frequent, and occur in every single DISCERN system following HLG. However, other agency shifts also occur that are not always as stable



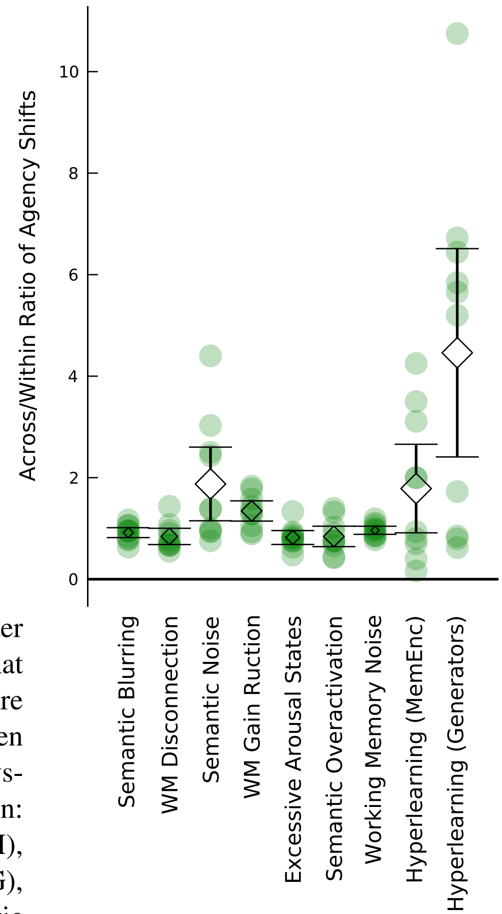


Figure 6.7: Agency shifts within the gangster context are compared with agency shifts that intrude from personal stories. Solid lines are averages as the lesion strengths increase; green circles are individual lesioned DISCERN systems. Four representative lesions are shown: hyperlearning of the memory encoder (HLM), hyperlearning of the generator modules (HLG), excessive arousal states (EA), and semantic noise (SN). HLG inserts personal characters into gangster stories more frequently than other lesions, suggesting that it impairs processing of context more.

and do not always produce new narratives that appear as meaningful. Agency shifts can also involve character from the same story context, which is not common in schizophrenia patients. One possible reason for this mismatch is that personal and gangster contexts in DISCERN are separated only implicitly by the presence of the self and other context-specific characters and content. In humans, real lived memories are qualitatively different from stories that are only experienced second-hand. Nevertheless, the data suggest that

DISCERN picks up on contextual cues, and that HLG impairs processing of context more (and other aspects of story processing less) than other lesions. Figure 6.7, for example, shows that HLG, more than other lesions, tends to insert personal characters into gangster stories, suggesting the insensitivity to context that is typical of schizophrenia.

6.6 Disorganized Speech

The HLG lesion (hyperlearning applied to the generator modules) was the focus of the findings presented so far. While it produces an interesting and plausible mechanism by which delusional stories and ideas may emerge, it does not produce many derailments, i.e. it does not tend to switch from one story to another in the middle of recall. This finding in itself is interesting, because derailments, just like delusions, are not shared by all patients with schizophrenia. However, a model of psychotic language in schizophrenia would not be complete if it did not account for mechanisms underlying derailed discourse.

Figure 6.8 illustrates the frequency of derailments for different lesions. The panel on the right shows the percentage of derailed language produced by all lesions at 40% recall. The two hyperlearning lesions differ dramatically in the amount of derailed language they produce: HLG causes very few derailments, while the language produced by HLM is derailed over 50% of the time on average, by far the most of all lesions. Several non-HL lesions, including SN and WMG, produce little derailed language, while others, including WMD and WMN, derail more frequently.

In the absence of patient data, the frequency of derailments alone cannot be used to decide that one lesion is a better model of language disorganization than another. A plausible model would be expected to produce a reasonable amount of derailed language, but more importantly, the language should not contain too many other errors that disrupt syntax and local story structure. Figure 6.9 shows that derailment-type errors dominate other errors for HLM but not for other lesions. The following example illustrates the way in which such word-level errors can make a story appear random and non-sensical rather than disorganized. It was produced by the WMN lesion (system #3, strength=0.2):

```
*Joe(I) went *St-Mary's(to).
I sat at a table.
I *drove(ordered) coffee.
*Tony(I) drank the coffee.
*Tony(I) met *Kate(Mary) at Starbucks.
[jumping to Story #12]
Joe was the fiancee of *Joe(Mary).
I hated *Bob(Joe).
I *feared(distrusted) *man(Joe).
I talked to Mary about wedding.
I *liked(hated) *the(wedding) *City-Hall(_).
*Kate(I) talked to *to(Mary) *the(a) *Four-Seasons(short).
I liked talking to *Joe(Mary).
I gave a *hand-shake(kiss) good-bye to *Vince(Mary).
```

DISCERN does derail to another story, but language like "I drove coffee" and "Kate talked to the Four Seasons" are too intrusive to create a credible simulation of derailed discourse. The output language of HLM, on the other hand, contains fewer grammatical and lexical errors, and when they do appear, they are much less disruptive:

```
[...]
I went to Four-Seasons.
I sat at a table.
I ordered wine.
I drank the wine.
I met *Stacy(Kate) at Four-Seasons.
[jumping to story #16]
Stacy was in her 20s.
Stacy had a ponytail.
Stacy was from New-York.
Stacy drove a compact car.
Stacy liked movies.
Stacy liked *baseball(books).
```

```
[jumping back to story #18]
I talked to Kate about *Kate(Stacy).
I liked *Mary(Stacy).
I talked to Kate a long time.
I liked talking to Kate.
I gave a kiss good-bye to *Mary(Kate).
```

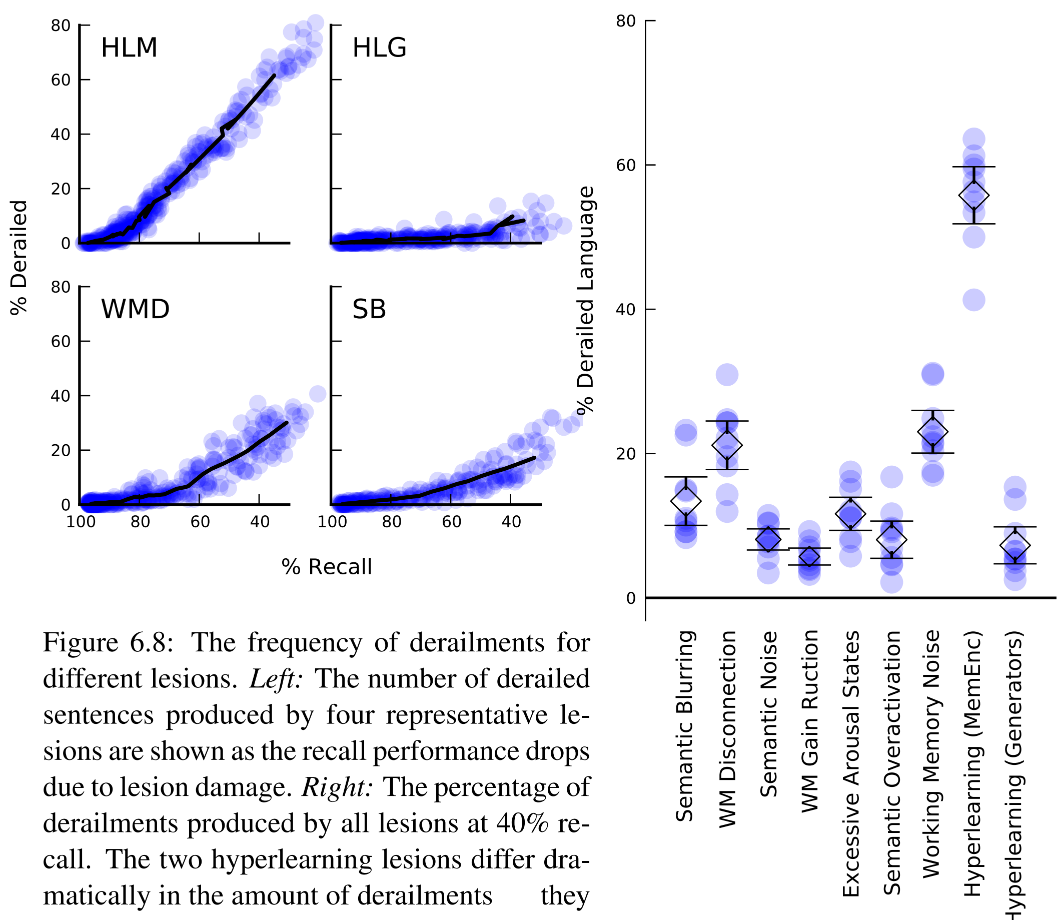


Figure 6.8: The frequency of derailments for different lesions. *Left:* The number of derailed sentences produced by four representative lesions are shown as the recall performance drops due to lesion damage. *Right:* The percentage of derailments produced by all lesions at 40% recall. The two hyperlearning lesions differ dramatically in the amount of derailments they produce: HLG derails only rarely, while HLM causes more derailments than any other lesion. Several non-HL lesions, including SN and WMG, produce little derailed language, while others, including WMD and WMN, derail more frequently.

This version of story #18 (system #2; HLM at strength = 1.5) is interesting for several other reasons as well. First, the jump to story 18 (which is about Stacy) is "foreshadowed" by the insertion of Stacy for Kate, which suggests that the jump is not merely an error in producing the correct memory cue, but that content from another story intrudes and interferes with DISCERN's discourse plan. Second, when DISCERN then jumps back to story #18, some confusion remains about who is talking to whom. Throughout the story, it appears as if DISCERN is trying to tell several similar stories at once. This impression is

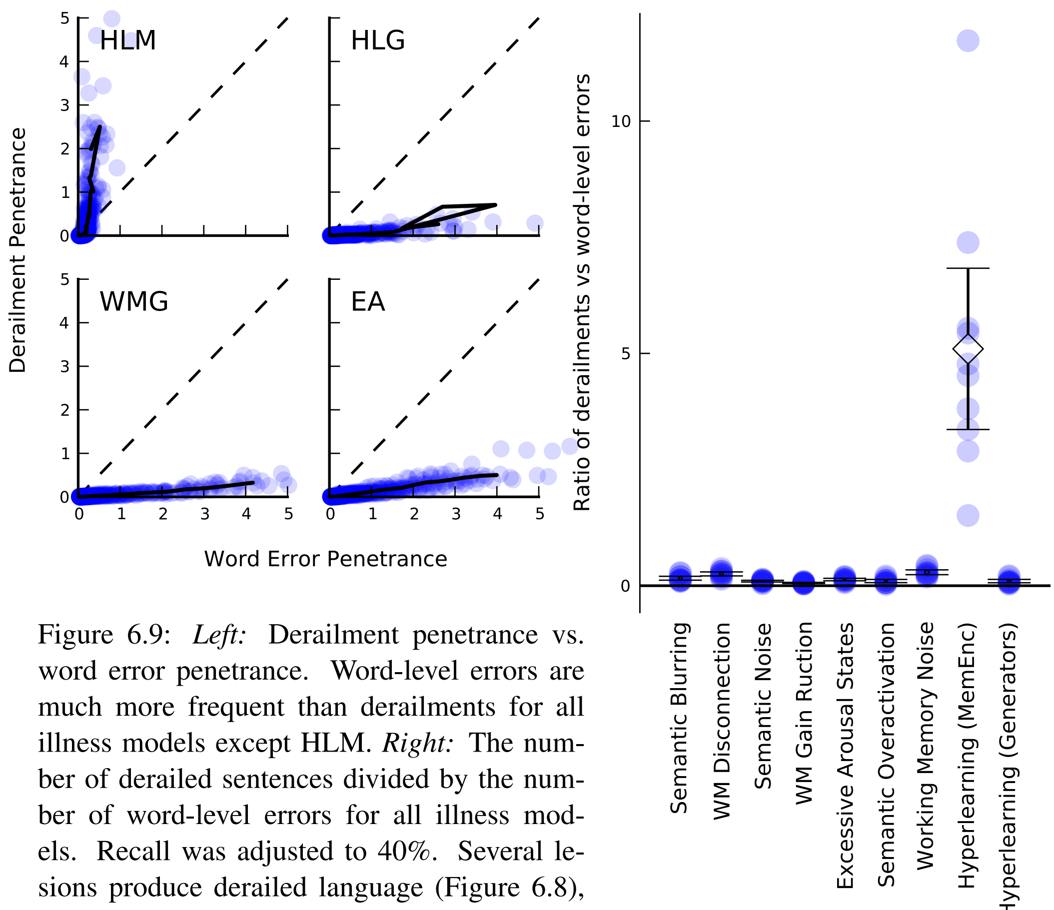


Figure 6.9: *Left:* Derailment penetrance vs. word error penetrance. Word-level errors are much more frequent than derailments for all illness models except HLM. *Right:* The number of derailed sentences divided by the number of word-level errors for all illness models. Recall was adjusted to 40%. Several lesions produce derailed language (Figure 6.8), but only HLM produces language where derailments dominate other errors.

even stronger in the similar example below, which was produced by a different DISCERN system (#4; HLM at 1.3) while recalling story #11:

```
[...]
Mary was my friend.
I loved Mary.
I trusted Mary.
[jumping to story #16]
I talked to *Mary(Stacy) about *guns(books).
I liked *baseball(books).
I talked to *Kate(Stacy) a long time.
I liked talking to *Mary(Stacy).
[jumping to story #18]
I talked to Kate about *Mary(Stacy).
I liked *Kate(Stacy).
I talked to Kate a long time.
I liked talking to *Mary(Kate).
I gave a kiss good-bye to *Mary(Kate).
```

Stories #11, #16, and #18 mix and interfere with each other, creating the impression of a fragmented narrative typical of language disorganization in schizophrenia. The same DISCERN system also produced interesting examples of the foreshadowing effect mentioned above. First, consider the following version of story # 19:

```
[...]
I wanted to go to the meeting.
I entered my car.
I drove to the Four-Seasons.
I was late.
I drove *carefully(recklessly).   [!]
[jumping to story #24]
Fred entered City-Hall for meeting.
Bob praised Fred.
```

```
The meeting was a success.
Fred gave a speech.
Fred drank champagne.
Fred ordered more champagne.
Fred got a-little drunk.
Fred had a good time.
```

Before derailing to story #24, "carefully" intrudes from that story to replace "recklessly". Stories #24 and #19 are similar, but when Fred drives to his meeting, he is on time and drives carefully, whereas I am late for my meeting, and drive recklessly. Later, when recalling story #24, the same system produced the following:

```
[...]
Fred wanted to go to City-Hall.
Fred entered his car.
Fred drove to City-Hall.
Fred was on-time.
Fred drove *recklessly(carefully).  [!]
Fred entered City-Hall for meeting.
[jumping to story #19]
I entered the Four-Seasons for meeting.
Joe praised me.
The meeting was a success.
I gave a speech.
I drank champagne.
[jumping back to story #24]
Fred ordered more champagne.
Fred got a-little drunk.
Fred had a good time.
```

This time, "recklessly" intrudes from story #19, foreshadowing an equivalent derailment in the other direction! Also note that the intrusion of "recklessly" is separated from the actual derailment by a sentence, so both were produced by separate cycles of the story generator.

Not all instances of derailed language produced by HLM were equally consistent with disorganized language in schizophrenia. In particular, at high lesion stregths, DISCERN tended to oscillate between two stories, as in the following example:

```
The Police investigated the bombing at City-Hall.
The Police looked for evidence.
The Police found that Tony bombed City-Hall.
The Police was after Tony.
[jumping to story #28]
The Police was after Vince.
The Police thought that Vince killed Bob.
The Police wanted to arrest(ed) Vince.
The Police found that Vince was in New-York.
The Police planned to arrest(ed) Vince in New-York.
[jumping to story #14]
I wanted to go to home.
I entered my car.
I drove to home.
I was drunk.
I drove recklessly.
[jumping to story #28]
Vince was pulled-over by a cop.
The cop asked Vince for his license.
[jumping to story #14]
I was pulled-over by a cop.
[jumping to story #28]
Vince was pulled-over by a cop.
The cop asked Vince for his license.
[jumping to story #14]
I was pulled-over by a cop.
```

```
[jumping to story #28]
Vince was pulled-over by a cop.
The cop asked Vince for his license.
[jumping to story #14]
I was pulled-over by a cop.
[...]
```

In this example, stories #28 and #14 intrude into the paraphrase of story #23. All three stories have similar content: A different character is pulled over by the police each time, and ends up getting arrested for a different reason. After the first jump to story #28, DISCERN begins to jump back and forth between #28 and #14 until recall is cut off for that story. Like before, the reason seems to be a conflict between two stories, resulting in a fragmented story. However, oscillations between stories are not common in schizophrenia, and were relatively frequent following HLM at high strength. Alternatively, the oscillations could be interpreted as perseveration, which is a fairly common sign of schizophrenic thought disorder (Andreasen 1979), and Maher et al. (1987) report that counted repetitions of words and phrases significantly correlate with rated levels of derailments in schizophrenia.

6.7 Conclusion

The experiments reported in this chapter attempted to reproduce the more intense symptoms that occur during active psychotic episodes in schizophrenia, focusing again on derailments and delusions. The main results were as follows. First, only the hyperlearning models were able to produce compelling simulations of psychosis in schizophrenia. When hyperlearning was applied to the generator modules of DISCERN, stable patterns of agency shifts ermerged that suggest a mechanism by which delusional narratives could be formed. This finding was unique to the HLG lesion – no other lesion produced similar patterns of agency shifts, and no other lesion showed the same tendency to insert the self across story contexts.

Applied to the memory encoder, hyperlearning led to frequent derailments but not to delusion-like language. Jumps to another story were often preceded by word insertions from that story, suggesting that they were caused by disturbances on a deeper level than that of faulty memory cues. Derailments often seemed to be caused by competition between stories, resulting in signs of fragmented discourse similar to that in schizophrenia.

One of the most intriguing findings of this experiment is that each version of hyperlearning produced a model for one psychotic symptom but not the other, depending on the part of the model to which it was applied. Since delusions and derailments are two hallmark symptoms of (respectively) paranoid-type and disorganized-type schizophrenia, this suggests that hyperlearing could model the emergence of clinical subtypes of schizophrenia from a shared underlying brain mechanism. This possibility is discussed further in the next chapter.

Chapter 7

Discussion and Future Work

The experiments reported above demonstrate that the DISCERN model can indeed be used to simulate and compare alternative illness mechanisms that could underlie schizophrenia. Candidate illness models can be characterized by the ways in which they distort storytelling in the model, and their viability and plausibility as models of schizophrenia can be judged by comparing the errors they cause to those made by schizophrenic patients. Creating a computational and experimental framework where hypotheses can be modeled and compared in this way was one of the main goals of this dissertation.

Beyond the proof of concept, a more ambitious goal was to create an illness model that actually captures important aspects of impaired story processing in schizophrenia, and that represents a plausible hypothesis about its causes. The hyperlearning mechanism located in different DISCERN modules produces such a hypothesis.

7.1 Summary of Results

The first set of experiments (Chapter 5) showed that hyperlearning can match the story-recall profile of human patients with schizophrenia (but not healthy controls) significantly better than other models. The resulting language of the best-fit hyperlearning models is also qualitatively similar to that produced by patients, including both derailments and delusion-like fixations. These findings suggest that hyperlearning captures specific aspects of pathophysiology underlying schizophrenia, rather than nonspecific sources of error-proneness demonstrated by human subjects overall.

In the second experimental study (Chapter 6), where the ability of the illness models to recreate more intense psychotic symptoms was tested, only the hyperlearning models were able to produce compelling simulations of psychotic language. Hyperlearning applied to DISCERN's generator modules caused stable patterns of agency shifts to ermerge in the output stories. Agents tended to cross over from autobiographical stories into gangster stories more than would be expected by chance, and more than was the case for other illness models. Hyperlearning also inserted the self into gangster stories very frequently, sometimes replacing every instance of a gangster in an entire story.

This tendency to confuse the self systematically with agents in gangster stories is one of the most interesting findings, because it creates a compelling model of a hallmark type of delusion in schizophrenia, i.e. the self-referential type. In these delusions, agents personally known to the patient (often the self) are confused with those in culturally endorsed narratives. Examples include a patient claiming that she is the Virgin Mary, and another that his upstairs neighbor is a CIA agent spying on him because he (the patient) has classified government information. The hyperlearning model suggests a mechanism by which such delusions could be formed: When the self crosses over into impersonal stories, spurious new narratives are created. It is easy to imagine that such stories, if they are remembered as real, could lead to complex and frightening delusional ideas. This model of fixed delusions was supported by data showing that patients with fixed narrative delusions made more agent-slotting errors than healthy controls and patients without these delusions. The model is unique to hyperlearning – no other illness model produced similar patterns of agency shifts, or had the same tendency to insert the self across story contexts.

When hyperlearning was applied to the memory encoder network, it caused frequent derailments but no delusion-like language. Jumps to another story were often preceded by word insertions from that story, suggesting that they were caused by disturbances of content rather than non-specific network error. Derailments often seemed to be caused by competition between stories, resulting in signs of fragmented discourse typical of schizophrenia. Furthermore, though several other illness models also caused derailments, only hyperlearning produced derailment behavior that was not accompanied by frequent lexical errors and break-down of syntax, which is not usually the case with real schizophrenic language.

One of the most intriguing findings of this research is that hyperlearning produced a model of delusions when applied to the generator modules, and of derailed speech when applied to the memory encoder. Since delusions and derailments are hallmark symptoms of (respectively) paranoid-type and disorganized-type schizophrenia, this suggests that hyperlearing could model how clinical subtypes of schizophrenia could share an underlying brain mechanism, but could emerge independently from each other. This explanation could also account for shared genetic vulnerabilities between subtypes, and for the fact that these subtypes are not necessarily stable over time. An encouraging preliminary results is that hyperlearning, applied to generators and memory encoder networks at the same time, can produce both delusional and derailed language in the same DISCERN system.

7.2 The Hyperlearning Hypothesis

The initial success of the hyperlearning model in simulating speakers with schizophrenia at different stages of the disorder suggests a promising direction for future research: Since hyperlearning was able to model both the behavior of patients in acute psychotic states (Chapter 6) and that of stable, medicated outpatients (Chapter 5), it seems likely that the transition from the former state to the latter could be modeled as well. If successful, the result would be the first simulation of antipsychotic drug action, potentially contributing to a better understanding of their effects, and possibly to the development of more effective treatments in the future. Modeling antipsychotic drug action as a learning process also explains why dopamine blockade does not relieve symptoms immediately.

The hyperlearning illness mechanism in the model arises from exaggerated backpropagation error signaling. Normal backpropagation in DISCERN and other connectionist models assumes a gradual consolidation process: Memories are replayed many times, and connection weights are adjusted in small, incremental steps (McClelland et al. 1995). Interestingly, normal memory consolidation in humans appears to involve repeated replay of memories as well (Euston et al. 2007), resulting in a gradual and incremental process that occurs over weeks or months (McGaugh 2000). Consequently, exaggerated backpropagation can plausibly simulate aberrant human memory consolidation due to overexuberant neuroplastic responses to prediction error (Kraus et al. 2009).

The hyperlearning hypothesis was inspired by Kapur's (2003) theory concerning motivational salience and psychosis, but it also converges with several recent behavioral and neurobiological studies. For instance, shared emotional valence in DISCERN plays an important role in triggering derailments, i.e. jumps from one story to another are very likely to involve two stories with the same or similar emotional valence. Similarly, emotionality has been found to prompt derailments in patients with schizophrenia (Docherty et al. 1998). Second, the narrative templates in DISCERN can be seen as components of social intelligence that predict goals and intentions of others (Bower and Morrow 1990); thus corrupted narrative memories modeled in DISCERN could account for impaired theory of mind in schizophrenia. Indeed delusions have been shown to be associated with an impaired capacity to understand the mental states of others (Bentall et al. 2009).

On the neurobiological side, in a recent imaging study healthy subjects were given a psychotomimetic drug (ketamine), and were then asked to perform an associative learning task. Greater cortical response to prediction error was associated with delusion formation when ketamine was administered (Corlett 2006). Furthermore, higher basal hippocampal/parahippocampal activity has been linked with schizophrenia, speech disorganization, and delusion formation (Moritz et al. 2003; Heckers et al. 1998; McGuire 1998; Schobel et al. 2009). These structures are central to memory consolidation (McGaugh 2000), and their increased activation may therefore indicate a hyperlearning mechanism. Memory consolidation can also be enhanced by elevated dopamine neurotransmission (Schott et al. 2006; Wittmann et al. 2005). Since schizophrenic psychosis has been linked to a hyperdopaminergic state (van Os and Kapur 2009), enhanced memory consolidation may well be the link between dopamine and psychosis.

The postulated hyperlearning mechanism should be studied experimentally in order to determine if it actually happens in psychotic patients. Specifically, the neural correlates of the predicted accelerated memory consolidation should be investigated. For instance, hyperlearning in humans could be observed via functional magnetic resonance imaging (fMRI; Takashima et al. 2009). This could be especially interesting during sleep, and in combination with behavioral measures of how fast and how well new memories are consolidated.

Another prediction of the hyperlearning model is that even brief periods of aberrant memory consolidation, perhaps lasting only days or weeks, could produce enduring psychosis. In DISCERN, a relatively small number of hyperlearning epochs (500) produced enduring memory reorganization. Preliminary results suggest that once this has happens, the "psychotic" network state is relatively stable, and can be reversed only with difficulty. This result could explain why delusions in patients often stay fixed over many years, and also why patients generally do not unlearn delusions but rather learn to limit their impact (Kapur 2003). This finding also suggests that hyperlearning in humans might be detectable as specific types of narrative learning impairments before the full syndrome of schizophrenia emerges (see for instance Brewer 2005), and novel interventions could be developed to mitigate its effects.

In addition to the predictions on the level of neurobiology, hyperlearning, as an explicit computational process, predicts specific changes of information processing in schizophrenia. The effects of hyperlearning in DISCERN should therefore be analyzed in detail, especially the way in which it affects the dynamics of story learning, and the way in which internal network representations and the flow of information in the model change. Understanding hyperlearning on this level promises further ways to test and validate the model, as well as a chance to gain new insights into the abnormal information processing that underlies psychosis in humans.

The hyperlearning mechanism is not only a plausible and predictive illness model, it also serves to demonstrate the strengths of the general modeling approach. First, hyperlearning was not part of the initial set of planned illness models. Instead, it emerged from preliminary experiments meant to explore the effects of compensatory network learning while other lesions were applied. In this way, computational models can suggest novel, alternative hypotheses through unexpected behavior. Second, several predictions of the model were similarly unexpected, and demonstrate how emergent behavior can suggest new explanations and tie together explanations of seemingly disparate symptoms. These predictions include delusion formation through agency shifts and the existence of a shared illness mechanism for different clinical subtypes. Third, hyperlearning demonstrates the conceptual reach of neural network modeling: The simulation is based on a relatively abstract theory, but offers a concrete, running interpretation of that theory, and is able to tie together research findings on different levels of analysis.

7.3 Extending the DISCERN Model Further

If hyperlearning occurs in humans, then the model is clearly a simplified version of the real process, just like DISCERN is a simplified version of human story processing. However, another advantage of computational models is that refinement and extensions are a natural part of their development. DISCERN was already extended significantly for this research, and can be extended further in future work. For example, one limitation of the current model is the absence of an explicit executive control mechanism, which could be added as a separate high-level control module (Miikkulainen 1993, 1996). Such a mechanism would make it possible to simulate executive dysfunction, which has been linked to both speech disorganization and delusions (Bentall et al. 2009; Kerns and Berenbaum 2003). The executive control model could include a model of attention and dopamine function, as suggested by (Braver et al. 1999). Motivational salience could be modeled within this framework as well, as a combination of emotional valence and unpredictability. DISCERN would then remember and generate only the salient parts of a story, which could make simulations of additional symptoms possible. For example a sign of thought disorder called *pressure of speech*, where patients talk rapidly and incessantly, could be simulated. Executive control could also provide an alternative model for alogia, simulating reduced output as a result of inattention and blunted emotions, and a refined version of hyperlearning, e.g. one where hyperlearning follows from network damage that increases prediction error.

Other possible extensions would put additional symptoms and illness mechanisms within reach. One straightforward example is the inclusion of pronouns and pronoun resolution. It is possible that this addition would only involve changes to the input stories, since DISCERN already uses possessive pronouns correctly, and there is no obvious reason why this ability would not scale up. Pronouns would make DISCERN's language significantly more realistic, and would also create an opportunity to model pronoun reference errors, a common kind of error in schizophrenia. Hallucinations are another important symptom that has so far not been modeled in DISCERN. Hallucinations cannot be observed directly through DISCERN's output language like derailments and delusional language. However, spurious speech perceptions in the parser modules could arise in response to network damage, which could be seen as a model of hallucinated speech. A previous similar approach using a speech perception SRN (Hoffman and McGlashan 1997, 2006) observed such spontaneous perceptions in the absence of external input.

As DISCERN and the illness models are refined further over time, a more complete picture of the mechanisms that are involved in the pathophysiology of schizophrenia may emerge. New experimental findings can be integrated into the model, possibly helping us integrate them into the overall understanding of schizophrenia. In the short term, the most important goal is to investigate experimentally whether the hypelearning mechanism happens in real psychotic patients. If validated, the hyperlearning hypothesis could contribute to a better understanding of schizophrenia, and provide a platform for characterizing and understanding the effects of current and future treatment interventions.

Chapter 8

Conclusion

Hallmark symptoms of schizophrenia are expressed as language behavior, specifically as defects of storytelling. These symptoms are diagnosed through clinical interviews, where narrative language is used as a window to the schizophrenic mind. This dissertation was motivated by the idea that computational models of schizophrenia should be able to do the same. Consequently, the research reported here was an attempt to understand the nature and pathophysiology of schizophrenia as disturbances in a computational model of story processing.

The main contribution of this dissertation is the first simulation of a speaker with schizophrenia. DISCERN, a neural network-based model of human story understanding and recall, was used to simulate how hypothetical neurobiological illness mechanisms could lead to abnormal storytelling observed in schizophrenia. The use of narrative language was the main feature that set this work apart from previous computational models of schizophrenia: In DISCERN, symptoms like delusions and derailments were observed directly at the level of narrative language — the same level at which real patients are diagnosed.

Based on the research literature, a range of candidate illness mechanisms were simulated in the model, and the resulting abnormal storytelling was evaluated and compared to that of patients with schizophrenia in two sets of computational experiments. First, data from a human subject study of story recall in schizophrenia was used to determine wich illness model was able to fit the story-recall profile of patients best. Second, the ability of the models to recreate the language of patients in acute psychotic phases of schizophrenia was evaluated, focusing on derailment behavior and signs of delusions.

Of all illness models, hyperlearning, a model of overly intense memory consolidation, produced the best fit to patient data, as well as compelling models of delusions and derailments. The hyperlearning/DISCERN model of language disturbance in schizophrenia is the second main contribution of this work. It represents a viable computational hypothesis about the way in which dopamine imbalance could lead to psychosis, and it ties together, and converges with, a range of previous research findings on dopamine imbalance, memory consolidation, and psychosis in schizophrenia.

Predictions of the hyperlearning model include the formation of fixed delusions through systematic confusion of agents and the existence of a shared underlying illness mechanism for different clinical subtypes of schizophrenia. Additionally, the hyperlearning hypothesis could be tested by studying the neural correlates of intensified memory consolidation in psychotic patients. If validated, the hyperlearning hypothesis could contribute to a better understanding of schizophrenia, and provide a platform for simulating the effects of medication and other future treatments.

> "I have transformed the problem from an intractably difficult and possibly quite insoluble conundrum into a mere linguistic puzzle. Albeit," he muttered, after a long moment of silent pondering, "an intractably difficult and possibly insoluble one."
>
> — Douglas Adams, Dirk Gently's Holistic Detective Agency

Appendix

Story Corpus

```
Story #1

Emotion: neutral      Context: Personal
[$person I 30s beard LA baseball beer rusty]
  I was in my 30s.                                    [I was 30s my _]
  I had a beard.                                    [I had _ _ beard]
  I was from LA.                                       [I was LA _ _]
  I drove a rusty car.                          [I drove _ rusty car]
  I liked baseball.                           [I liked _ _ baseball]
  I liked beer.                                   [I liked _ _ beer]
[$job I St-Vincent's worked liked New-York good doctor]
  I was a doctor.                                 [I was _ _ doctor]
  I worked for St-Vincent's.            [I worked St-Vincent's _ _]
  I worked in New-York.                     [I worked New-York _ _]
  I liked my job.                                 [I liked _ my job]
  I was a good doctor.                         [I was _ good doctor]
story-end

Story #2

Emotion: very negative      Context: Personal
[$person Joe 30s beard Chicago baseball wine nice]
  Joe was in his 30s.                            [Joe 2was 30s his _]
  Joe had a beard.                                [Joe had _ _ beard]
  Joe was from Chicago.                        [Joe was Chicago _ _]
```

```
  Joe drove a nice car.                        [Joe drove _ nice car]
  Joe liked baseball.                        [Joe liked _ _ baseball]
  Joe liked wine.                                [Joe liked _ _ wine]
[$job Joe St-Vincent's head liked New-York famous doctor]
  Joe was a doctor.                             [Joe was _ _ doctor]
  Joe was the head of St-Vincent's.     [Joe was St-Vincent's _ head]
  Joe worked in New-York.                  [Joe worked New-York _ _]
  Joe liked his job.                           [Joe liked _ his job]
  Joe was a famous doctor.                 [Joe was _ famous doctor]
[$relation I Joe hated distrusted _ my boss]
  Joe was my boss.                               [Joe was _ my boss]
  I hated Joe.                                     [I hated _ _ Joe]
  I distrusted Joe.                            [I distrusted _ _ Joe]
story-end

Story #3

Emotion: very positive      Context: Personal
[$person Mary 30s ponytail LA movies books compact]
  Mary was in her 30s.                          [Mary was 30s her _]
  Mary had a ponytail.                      [Mary had _ _ ponytail]
  Mary was from LA.                               [Mary was LA _ _]
  Mary drove a compact car.              [Mary drove _ compact car]
  Mary liked movies.                        [Mary liked _ _ movies]
  Mary liked books.                          [Mary liked _ _ books]
[$relation I Mary loved trusted _ my friend]
  Mary was my friend.                        [Mary was _ my friend]
  I loved Mary.                                   [I loved _ _ Mary]
  I trusted Mary.                               [I trusted _ _ Mary]
story-end


Story #4

Emotion: very positive      Context: Personal
[$relation I Stacy liked trusted _ my girlfriend]
  Stacy was my girlfriend.              [Stacy was _ my girlfriend]
  I liked Stacy.                                  [I liked _ _ Stacy]
  I trusted Stacy.                              [I trusted _ _ Stacy]
```

```
[$person Stacy 20s ponytail New-York baseball New-York compact]
  Stacy was in her 20s.                         [Stacy was 20s her _]
  Stacy had a ponytail.                      [Stacy had _ _ ponytail]
  Stacy was from New-York.                   [Stacy was New-York _ _]
  Stacy drove a compact car.              [Stacy drove _ compact car]
  Stacy liked baseball.                    [Stacy liked _ _ baseball]
  Stacy liked New-York.                    [Stacy liked _ _ New-York]
story-end


Story #5

Emotion: positive      Context: Personal
[$person Kate 50s ponytail LA books wine nice]
  Kate was in her 50s.                           [Kate was 50s her _]
  Kate had a ponytail.                        [Kate had _ _ ponytail]
  Kate was from LA.                                  [Kate was LA _ _]
  Kate drove a nice car.                      [Kate drove _ nice car]
  Kate liked books.                            [Kate liked _ _ books]
  Kate liked wine.                              [Kate liked _ _ wine]
[$relation I Kate loved trusted _ my mother]
  Kate was my mother.                          [Kate was _ my mother]
  I loved Kate.                                    [I loved _ _ Kate]
  I trusted Kate.                                [I trusted _ _ Kate]
story-end


Story #6

Emotion: neutral      Context: Gangster
[$person Fred 30s mustache New-York baseball guns rusty]
  Fred was in his 30s.                           [Fred was 30s his _]
  Fred had a mustache.                        [Fred had _ _ mustache]
  Fred was from New-York.                     [Fred was New-York _ _]
  Fred drove a rusty car.                    [Fred drove _ rusty car]
  Fred liked baseball.                      [Fred liked _ _ baseball]
  Fred liked guns.                              [Fred liked _ _ guns]
```

```
[$job Fred Police worked liked New-York good cop]
  Fred was a cop.                                  [Fred was _ _ cop]
  Fred worked for the Police.                [Fred worked Police _ _]
  Fred worked in New-York.                 [Fred worked New-York _ _]
  Fred liked his job.                          [Fred liked _ his job]
  Fred was a good cop.                          [Fred was _ good cop]
story-end
```

Story #7

```
Emotion: neutral      Context: Gangster
[$person Bob 50s mustache Chicago guns baseball nice]
  Bob was in his 50s.                              [Bob was 50s his _]
  Bob had a mustache.                          [Bob had _ _ mustache]
  Bob was from Chicago.                         [Bob was Chicago _ _]
  Bob drove a nice car.                        [Bob drove _ nice car]
  Bob liked guns.                                [Bob liked _ _ guns]
  Bob liked baseball.                        [Bob liked _ _ baseball]
[$job Bob Police head liked City-Hall good cop]
  Bob was a cop.                                    [Bob was _ _ cop]
  Bob was the head of the Police.             [Bob was Police _ head]
  Bob worked at City-Hall.                 [Bob worked City-Hall _ _]
  Bob liked his job.                            [Bob liked _ his job]
  Bob was a good cop.                            [Bob was _ good cop]
[$relation Fred Bob liked trusted _ Fred boss]
  Bob was the boss of Fred.                     [Bob was Fred _ boss]
  Fred liked Bob.                                [Fred liked _ _ Bob]
  Fred trusted Bob.                            [Fred trusted _ _ Bob]
story-end
```

Story #8

```
Emotion: neutral      Context: Gangster
[$person Tony 20s mustache Chicago baseball guns rusty]
  Tony was in his 20s.                            [Tony was 20s his _]
  Tony had a mustache.                        [Tony had _ _ mustache]
  Tony was from Chicago.                       [Tony was Chicago _ _]
  Tony drove a rusty car.                    [Tony drove _ rusty car]
```

```
  Tony liked baseball.                     [Tony liked _ _ baseball]
  Tony liked guns.                             [Tony liked _ _ guns]
[$job Tony Mafia worked hated Chicago bad gangster]
  Tony was a gangster.                       [Tony was _ _ gangster]
  Tony worked for the Mafia.                 [Tony worked Mafia _ _]
  Tony worked in Chicago.                  [Tony worked Chicago _ _]
  Tony hated his job.                         [Tony hated _ his job]
  Tony was a bad gangster.                 [Tony was _ bad gangster]
[$relation Vince Tony hated distrusted _ Vince co-worker]
  Tony was a co-worker of Vince.        [Tony was Vince _ co-worker]
  Vince hated Tony.                           [Vince hated _ _ Tony]
  Vince distrusted Tony.                 [Vince distrusted _ _ Tony]
story-end


Story #9

Emotion: neutral      Context: Gangster
[$person Vince 30s beard LA guns movies nice]
  Vince was in his 30s.                        [Vince was 30s his _]
  Vince had a beard.                           [Vince had _ _ beard]
  Vince was from LA.                              [Vince was LA _ _]
  Vince drove a nice car.                   [Vince drove _ nice car]
  Vince liked guns.                           [Vince liked _ _ guns]
  Vince liked movies.                       [Vince liked _ _ movies]
[$job Vince Mafia worked liked LA good gangster]
  Vince was a gangster.                     [Vince was _ _ gangster]
  Vince worked for the Mafia.               [Vince worked Mafia _ _]
  Vince worked in LA.                          [Vince worked LA _ _]
  Vince liked his job.                       [Vince liked _ his job]
  Vince was a good gangster.             [Vince was _ good gangster]
[$relation Tony Vince hated feared _ Tony co-worker]
  Vince was a co-worker of Tony.        [Vince was Tony _ co-worker]
  Tony hated Vince.                           [Tony hated _ _ Vince]
  Tony feared Vince.                         [Tony feared _ _ Vince]
story-end
```

Story #10

```
Emotion: neutral      Context: Gangster
[$person Vito 50s beard New-York guns baseball nice]
  Vito was in his 50s.                          [Vito was 50s his _]
  Vito had a beard.                             [Vito had _ _ beard]
  Vito was from New-York.                      [Vito was New-York _ _]
  Vito drove a nice car.                       [Vito drove _ nice car]
  Vito liked guns.                               [Vito liked _ _ guns]
  Vito liked baseball.                       [Vito liked _ _ baseball]
[$job Vito Mafia head liked New-York famous gangster]
  Vito was a gangster.                         [Vito was _ _ gangster]
  Vito was the head of the Mafia.              [Vito was Mafia _ head]
  Vito worked in New-York.                  [Vito worked New-York _ _]
  Vito liked his job.                         [Vito liked _ his job]
  Vito was a famous gangster.            [Vito was _ famous gangster]
[$relation Vince Vito liked feared _ Vince boss]
  Vito was the boss of Vince.                  [Vito was Vince _ boss]
  Vince liked Vito.                           [Vince liked _ _ Vito]
  Vince feared Vito.                         [Vince feared _ _ Vito]
[$relation Tony Vito hated feared _ Tony boss]
  Vito was the boss of Tony.                    [Vito was Tony _ boss]
  Tony hated Vito.                              [Tony hated _ _ Vito]
  Tony feared Vito.                            [Tony feared _ _ Vito]
story-end
```

Story #11

```
Emotion: very positive      Context: Personal
[$drink I Mary met Moe's-Tavern table wine _]
  I went to Moe's-Tavern.                   [I went Moe's-Tavern _ _]
  I sat at a table.                                [I sat table _ _]
  I ordered wine.                                [I ordered _ _ wine]
  I drank the wine.                                [I drank _ _ wine]
  I met Mary at Moe's-Tavern.            [I met Moe's-Tavern _ Mary]
[$person Mary 20s ponytail LA movies books compact]
  Mary was in her 20s.                          [Mary was 20s her _]
  Mary had a ponytail.                       [Mary had _ _ ponytail]
```

```
  Mary was from LA.                              [Mary was LA _ _]
  Mary drove a compact car.                [Mary drove _ compact car]
  Mary liked movies.                          [Mary liked _ _ movies]
  Mary liked books.                            [Mary liked _ _ books]
[$relation I Mary loved trusted _ my friend]
  Mary was my friend.                          [Mary was _ my friend]
  I loved Mary.                                    [I loved _ _ Mary]
  I trusted Mary.                                [I trusted _ _ Mary]
[$talking I Mary liked liked kiss movies long]
  I talked to Mary about movies.         [I talked Mary about movies]
  I liked movies.                                [I liked _ _ movies]
  I talked to Mary a long time.             [I talked Mary long time]
  I liked talking to Mary.                   [I liked Mary _ talking]
  I gave a kiss good-bye to Mary.         [I gave Mary kiss good-bye]
story-end
```

Story #12

```
Emotion: negative      Context: Personal
[$drink I Joe met Moe's-Tavern counter beer _]
  I went to Moe's-Tavern.                   [I went Moe's-Tavern _ _]
  I sat at the counter.                         [I sat counter _ _]
  I ordered beer.                                [I ordered _ _ beer]
  I drank the beer.                                [I drank _ _ beer]
  I met Joe at Moe's-Tavern.               [I met Moe's-Tavern _ Joe]
[$relation I Joe hated distrusted _ my boss]
  Joe was my boss.                                [Joe was _ my boss]
  I hated Joe.                                      [I hated _ _ Joe]
  I distrusted Joe.                            [I distrusted _ _ Joe]
[$person Joe 30s beard Chicago baseball wine nice]
  Joe was in his 30s.                             [Joe was 30s his _]
  Joe had a beard.                                [Joe had _ _ beard]
  Joe was from Chicago.                         [Joe was Chicago _ _]
  Joe drove a nice car.                        [Joe drove _ nice car]
  Joe liked baseball.                        [Joe liked _ _ baseball]
  Joe liked wine.                                [Joe liked _ _ wine]
```

```
[$relation Mary Joe hated distrusted _ I fiancee]
  Joe was the fiancee of Mary.              [Joe was Mary _ fiancee]
  I hated Joe.                                     [I hated _ _ Joe]
  I distrusted Joe.                           [I distrusted _ _ Joe]
[$talking I Joe loved hated hand-shake Mary short]
  I talked to Joe about Mary.              [I talked Joe about Mary]
  I loved Mary.                                   [I loved _ _ Mary]
  I talked to Joe a short time.            [I talked Joe short time]
  I hated talking to Joe.                   [I hated Joe _ talking]
  I gave a hand-shake good-bye to Joe.
                                     [I gave Joe hand-shake good-bye]
story-end
```

Story #13

```
Emotion: negative      Context: Personal
[$drink I Mary met Starbucks table coffee _]
  I went to Starbucks.                        [I went Starbucks _ _]
  I sat at a table.                                [I sat table _ _]
  I ordered coffee.                           [I ordered _ _ coffee]
  I drank the coffee.                           [I drank _ _ coffee]
  I met Mary at Starbucks.                  [I met Starbucks _ Mary]
[$relation Joe Mary loved trusted _ I fiancee]
  Mary was the fiancee of Joe.              [Mary was Joe _ fiancee]
  I loved Mary.                                   [I loved _ _ Mary]
  I trusted Mary.                               [I trusted _ _ Mary]
[$talking I Mary hated liked kiss wedding short]
  I talked to Mary about wedding.      [I talked Mary about wedding]
  I hated wedding.                              [I hated _ _ wedding]
  I talked to Mary a short time.          [I talked Mary short time]
  I liked talking to Mary.                  [I liked Mary _ talking]
  I gave a kiss good-bye to Mary.        [I gave Mary kiss good-bye]
story-end
```

```
Story #14

Emotion: very negative      Context: Personal
[$drink I man met Moe's-Tavern counter beer _]
  I went to Moe's-Tavern.                    [I went Moe's-Tavern _ _]
  I sat at the counter.                            [I sat counter _ _]
  I ordered beer.                                 [I ordered _ _ beer]
  I drank the beer.                                 [I drank _ _ beer]
  I met man at Moe's-Tavern.                [I met Moe's-Tavern _ man]
[$talking I man loved hated hand-shake Mary long]
  I talked to man about Mary.                [I talked man about Mary]
  I loved Mary.                                     [I loved _ _ Mary]
  I talked to man a long time.                [I talked man long time]
  I hated talking to man.                      [I hated man _ talking]
  I gave a hand-shake good-bye to man.
                                     [I gave man hand-shake good-bye]
[$drunk I _ _ _ beer bad very]
  I ordered more beer.                       [I ordered _ more beer]
  I got very drunk.                             [I got _ very drunk]
  I had a bad time.                               [I had _ bad time]
[$driving I _ drunk home home recklessly _]
  I wanted to go to home.                       [I wanted home go _]
  I entered my car.                             [I entered _ my car]
  I drove to home.                                [I drove home _ _]
  I was drunk.                                     [I was _ _ drunk]
  I drove recklessly.                     [I drove _ _ recklessly]
[$pulled-over I cop arrest(ed) _ DUI _ _]
  I was pulled-over by a cop.             [I was cop _ pulled-over]
  The cop asked me for my license.         [cop asked license my I]
  I gave my license to The cop.            [I gave cop my license]
  The cop checked the license.            [cop checked _ _ license]
  The cop arrest(ed) me for DUI.          [cop arrest(ed) DUI _ I]
[$trial I _ got convicted fine DUI bad]
  I was accused of DUI.                      [I was DUI _ accused]
  I was brought before the court.          [I was court _ brought]
  I had a bad lawyer.                          [I had _ bad lawyer]
  The court convicted me of DUI.          [court convicted DUI _ I]
  I got a fine.                                    [I got _ fine _]
story-end
```

Story #15

```
Emotion: negative      Context: Personal
[$drink Joe Mary met Four-Seasons table champagne _]
  Joe went to Four-Seasons.                [Joe went Four-Seasons _ _]
  Joe sat at a table.                              [Joe sat table _ _]
  Joe ordered champagne.                   [Joe ordered _ _ champagne]
  Joe drank the champagne.                   [Joe drank _ _ champagne]
  Joe met Mary at Four-Seasons.          [Joe met Four-Seasons _ Mary]
[$relation Joe Mary loved trusted _ Joe girlfriend]
  Mary was the girlfriend of Joe.          [Mary was Joe _ girlfriend]
  Joe loved Mary.                                 [Joe loved _ _ Mary]
  Joe trusted Mary.                             [Joe trusted _ _ Mary]
[$talking Joe Mary liked liked kiss wedding long]
  Joe talked to Mary about wedding.   [Joe talked Mary about wedding]
  Joe liked the wedding.                        [Joe liked _ _ wedding]
  Joe talked to Mary a long time.          [Joe talked Mary long time]
  Joe liked talking to Mary.                [Joe liked Mary _ talking]
  Joe gave a kiss good-bye to Mary.     [Joe gave Mary kiss good-bye]
[$plan Mary people invite LA wedding I invitation]
  Mary planned a wedding in LA.            [Mary planned LA _ wedding]
  Mary wanted to invite a-lot-of people.[Mary wanted _ invite people]
  Mary sent invitation to me.               [Mary sent I _ invitation]
  I accepted the invitation.               [I accepted _ _ invitation]
[$drunk Joe _ _ _ champagne good a-little]
  Joe ordered more champagne.           [Joe ordered _ more champagne]
  Joe got a-little drunk.                   [Joe got _ a-little drunk]
  Joe had a good time.                           [Joe had _ good time]
story-end
```

Story #16

```
Emotion: positive      Context: Personal
[$drink I Stacy met Moe's-Tavern table wine _]
  I went to Moe's-Tavern.                    [I went Moe's-Tavern _ _]
  I sat at a table.                                  [I sat table _ _]
  I ordered wine.                                 [I ordered _ _ wine]
  I drank the wine.                                 [I drank _ _ wine]
  I met Stacy at Moe's-Tavern.           [I met Moe's-Tavern _ Stacy]
```

```
[$relation I Stacy liked trusted _ my girlfriend]
  Stacy was my girlfriend.                [Stacy was _ my girlfriend]
  I liked Stacy.                                  [I liked _ _ Stacy]
  I trusted Stacy.                              [I trusted _ _ Stacy]
[$person Stacy 20s ponytail New-York movies books compact]
  Stacy was in her 20s.                         [Stacy was 20s her _]
  Stacy had a ponytail.                      [Stacy had _ _ ponytail]
  Stacy was from New-York.                   [Stacy was New-York _ _]
  Stacy drove a compact car.              [Stacy drove _ compact car]
  Stacy liked movies.                        [Stacy liked _ _ movies]
  Stacy liked books.                          [Stacy liked _ _ books]
[$talking I Stacy  liked liked kiss books long]
  I talked to Stacy about books.         [I talked Stacy about books]
  I liked books.                                  [I liked _ _ books]
  I talked to Stacy a long time.           [I talked Stacy long time]
  I liked talking to Stacy.                 [I liked Stacy _ talking]
  I gave a kiss good-bye to Stacy.       [I gave Stacy kiss good-bye]
story-end
```

```
Story #17

Emotion: very negative      Context: Personal
[$driving I _ late Four-Seasons wedding carefully _]
  I wanted to go to wedding.                  [I wanted wedding go _]
  I entered my car.                              [I entered _ my car]
  I drove to the Four-Seasons.             [I drove Four-Seasons _ _]
  I was late.                                        [I was _ _ late]
  I drove carefully.                          [I drove _ _ carefully]
[$occasion I Mary speech kissed Four-Seasons Joe wedding]
  I entered the Four-Seasons for wedding.
                                    [I entered wedding _ Four-Seasons]
  Mary kissed Joe.                              [Mary kissed _ _ Joe]
  The wedding was a success.               [wedding was _ _ success]
  I gave a speech.                               [I gave _ _  speech]
  I drank champagne.                          [I drank _ _ champagne]
[$talking I Mary hated hated kiss wedding short]
  I talked to Mary about wedding.       [I talked Mary about wedding]
  I hated wedding.                              [I hated _ _ wedding]
```

```
  I talked to Mary a short time.          [I talked Mary short time]
  I hated talking to Mary.                  [I hated Mary _ talking]
  I gave a kiss good-bye to Mary.         [I gave Mary kiss good-bye]
[$drunk I _ _ _ champagne bad very]
  I ordered more champagne.             [I ordered _ more champagne]
  I got very drunk.                              [I got _ very drunk]
  I had a bad time.                                [I had _ bad time]
story-end
```

Story #18

```
Emotion: positive      Context: Personal
[$plan I Kate met LA meeting Kate invitation]
  I planned a meeting in LA.                [I planned LA _ meeting]
  I wanted to met Kate.                        [I wanted _ met Kate]
  I sent invitation to Kate.              [I sent Kate _ invitation]
  Kate accepted the invitation.        [Kate accepted _ _ invitation]
[$relation I Kate loved trusted _ my mother]
  Kate was my mother.                         [Kate was _ my mother]
  I loved Kate.                                    [I loved _ _ Kate]
  I trusted Kate.                                [I trusted _ _ Kate]
[$flight I no-one _ New-York LA long late]
  I entered the New-York airport.      [I entered _ New-York airport]
  I walked to the counter.                    [I walked counter _ _]
  I checked my bag.                              [I checked _ my bag]
  I walked to my gate.                          [I walked gate my _]
  I entered the plane to LA.                  [I entered LA _ plane]
  The plane was late.                            [plane was _ _ late]
  I met no-one in the plane.                  [I met plane _ no-one]
[$drink I Kate met Four-Seasons table wine _]
  I went to Four-Seasons.                [I went Four-Seasons _ _]
  I sat at a table.                                 [I sat table _ _]
  I ordered wine.                                [I ordered _ _ wine]
  I drank the wine.                                [I drank _ _ wine]
  I met Kate at Four-Seasons.            [I met Four-Seasons _ Kate]
[$person Kate 50s ponytail LA books wine nice]
  Kate was in her 50s.                           [Kate was 50s her _]
  Kate had a ponytail.                        [Kate had _ _ ponytail]
```

Kate was from LA. [Kate was LA _ _]
Kate drove a nice car. [Kate drove _ nice car]
Kate liked books. [Kate liked _ _ books]
Kate liked wine. [Kate liked _ _ wine]
[$talking I Kate liked liked kiss Stacy long]
I talked to Kate about Stacy. [I talked Kate about Stacy]
I liked Stacy. [I liked _ _ Stacy]
I talked to Kate a long time. [I talked Kate long time]
I liked talking to Kate. [I liked Kate _ talking]
I gave a kiss good-bye to Kate. [I gave Kate kiss good-bye]
story-end

Story #19

Emotion: positive Context: Personal
[$job Joe St-Vincent’s head liked New-York famous doctor]
Joe was a doctor. [Joe was _ _ doctor]
Joe was the head of St-Vincent’s. [Joe was St-Vincent’s _ head]
Joe worked in New-York. [Joe worked New-York _ _]
Joe liked his job. [Joe liked _ his job]
Joe was a famous doctor. [Joe was _ famous doctor]
[$relation I Joe hated distrusted _ my boss]
Joe was my boss. [Joe was _ my boss]
I hated Joe. [I hated _ _ Joe]
I distrusted Joe. [I distrusted _ _ Joe]
[$plan Joe people praise(d) New-York Meeting I invitation]
Joe planned a Meeting in New-York. [Joe planned New-York _ Meeting]
Joe wanted to praise(d) a-lot-of people.
[Joe wanted _ praise(d) people]
Joe sent invitation to me. [Joe sent I _ invitation]
I accepted the invitation. [I accepted _ _ invitation]
[$driving I _ late Four-Seasons meeting recklessly _]
I wanted to go to the meeting. [I wanted meeting go _]
I entered my car. [I entered _ my car]
I drove to the Four-Seasons. [I drove Four-Seasons _ _]
I was late. [I was _ _ late]
I drove recklessly. [I drove _ _ recklessly]

```
[$occasion I Joe speech praise(d) Four-Seasons me meeting]
  I entered the Four-Seasons for meeting.
                                   [I entered meeting _ Four-Seasons]
  Joe praise(d) me.                            [Joe praise(d) _ _ me]
  The meeting was a success.                [meeting was _ _ success]
  I gave a speech.                               [I gave _ _  speech]
  I drank champagne.                         [I drank _ _ champagne]
[$drunk I _ _ _ champagne good a-little]
  I ordered more champagne.              [I ordered _ more champagne]
  I got a-little drunk.                      [I got _ a-little drunk]
  I had a good time.                              [I had _ good time]
story-end

Story #20

Emotion: very negative      Context: Gangster
[$job Vito Mafia head liked New-York famous gangster]
  Vito was a gangster.                       [Vito was _ _ gangster]
  Vito was the head of the Mafia.            [Vito was Mafia _ head]
  Vito worked in New-York.                [Vito worked New-York _ _]
  Vito liked his job.                         [Vito liked _ his job]
  Vito was a famous gangster.          [Vito was _ famous gangster]
[$plan Vito Tony met New-York meeting Tony invitation]
  Vito planned a meeting in New-York.  [Vito planned New-York _ mtg.]
  Vito wanted to met Tony.                  [Vito wanted _ met Tony]
  Vito sent invitation to Tony.        [Vito sent Tony _ invitation]
  Tony accepted the invitation.       [Tony accepted _ _ invitation]
[$relation Tony Vito hated feared _ Tony boss]
  Vito was the boss of Tony.                  [Vito was Tony _ boss]
  Tony hated Vito.                             [Tony hated _ _ Vito]
  Tony feared Vito.                           [Tony feared _ _ Vito]
[$flight Tony no-one _ Chicago New-York long late]
  Tony entered the Chicago airport.  [Tony entered _ Chicago airport]
  Tony walked to the counter.            [Tony walked counter _ _]
  Tony checked his bag.                      [Tony checked _ his bag]
  Tony walked to his gate.                  [Tony walked gate his _]
  Tony entered the plane to New-York. [Tony entered New-York _ plane]
  The plane was late.                            [plane was _ _ late]
  Tony met no-one in the plane.            [Tony met plane _ no-one]
```

```
[$drink Tony Vito met Four-Seasons table wine _]
  Tony went to Four-Seasons.            [Tony went Four-Seasons _ _]
  Tony sat at a table.                          [Tony sat table _ _]
  Tony ordered wine.                         [Tony ordered _ _ wine]
  Tony drank the wine.                         [Tony drank _ _ wine]
  Tony met Vito at Four-Seasons.      [Tony met Four-Seasons _ Vito]
[$plan Vito City-Hall bomb(ed) New-York bombing Tony order]
  Vito planned a bombing in New-York.
                                  [Vito planned New-York _ bombing]
  Vito wanted to bomb(ed) City-Hall.
                                 [Vito wanted _ bomb(ed) City-Hall]
  Vito gave order to Tony.                  [Vito gave Tony _ order]
  Tony accepted the order.                 [Tony accepted _ _ order]
story-end
```

Story #21

```
Emotion: very negative      Context: Gangster
[$job Tony Mafia worked hated New-York bad gangster]
  Tony was a gangster.                       [Tony was _ _ gangster]
  Tony worked for the Mafia.                 [Tony worked Mafia _ _]
  Tony worked in New-York.                [Tony worked New-York _ _]
  Tony hated his job.                          [Tony hated _ his job]
  Tony was a bad gangster.                 [Tony was _ bad gangster]
[$driving Tony _ scared City-Hall City-Hall carefully _]
  Tony wanted to go to City-Hall.        [Tony wanted City-Hall go _]
  Tony entered his car.                     [Tony entered _ his car]
  Tony drove to City-Hall.                [Tony drove City-Hall _ _]
  Tony was scared.                             [Tony was _ _ scared]
  Tony drove carefully.                 [Tony drove _ _ carefully]
[$occasion Tony Tony phone-call bomb(ed) City-Hall City-Hall bombing]
  Tony entered City-Hall for bombing.
                                 [Tony entered bombing _ City-Hall]
  Tony bomb(ed) City-Hall.              [Tony bomb(ed) _ _ City-Hall]
  The bombing was a success.              [bombing was _ _ success]
  Tony made a phone-call.               [Tony made _ _  phone-call]
  Tony smoked a cigarette.             [Tony smoked _ _ cigarette]
```

```
[$drink Tony no-one met Moe's-Tavern counter beer _]
  Tony went to Moe's-Tavern.             [Tony went Moe's-Tavern _ _]
  Tony sat at the counter.                     [Tony sat counter _ _]
  Tony ordered beer.                          [Tony ordered _ _ beer]
  Tony drank the beer.                          [Tony drank _ _ beer]
  Tony met no-one at Moe's-Tavern.   [Tony met Moe's-Tavern _ no-one]
[$drunk Tony _ _ _ beer bad very]
  Tony ordered more beer.                 [Tony ordered _ more beer]
  Tony got very drunk.                       [Tony got _ very drunk]
  Tony had a bad time.                         [Tony had _ bad time]
story-end


Story #22

Emotion: neutral      Context: Gangster
[$drink Vince Vito met Moe's-Tavern counter beer _]
  Vince went to Moe's-Tavern.           [Vince went Moe's-Tavern _ _]
  Vince sat at the counter.                   [Vince sat counter _ _]
  Vince ordered beer.                        [Vince ordered _ _ beer]
  Vince drank the beer.                        [Vince drank _ _ beer]
  Vince met Vito at Moe's-Tavern.     [Vince met Moe's-Tavern _ Vito]
[$relation Vince Vito liked feared _ Vince boss]
  Vito was the boss of Vince.                 [Vito was Vince _ boss]
  Vince liked Vito.                           [Vince liked _ _ Vito]
  Vince feared Vito.                         [Vince feared _ _ Vito]
[$person Vito 30s beard Chicago guns movies nice]
  Vito was in his 30s.                           [Vito was 30s his _]
  Vito had a beard.                             [Vito had _ _ beard]
  Vito was from Chicago.                     [Vito was Chicago _ _]
  Vito drove a nice car.                     [Vito drove _ nice car]
  Vito liked guns.                             [Vito liked _ _ guns]
  Vito liked movies.                         [Vito liked _ _ movies]
[$talking Vince Vito liked liked kiss guns long]
  Vince talked to Vito about guns.     [Vince talked Vito about guns]
  Vince liked guns.                           [Vince liked _ _ guns]
  Vince talked to Vito a long time.     [Vince talked Vito long time]
  Vince liked talking to Vito.           [Vince liked Vito _ talking]
  Vince gave a kiss good-bye to Vito. [Vince gave Vito kiss good-bye]
```

```
[$drunk Vince _ _ _ beer good a-little]
  Vince ordered more beer.              [Vince ordered _ more beer]
  Vince got a-little drunk.            [Vince got _ a-little drunk]
  Vince had a good time.                    [Vince had _ good time]
story-end


Story #23

Emotion: negative      Context: Gangster
[$investigation Tony Police bomb(ed) City-Hall City-Hall bombing _]
  The Police investigated the bombing at City-Hall.
                            [Police investigated City-Hall _ bombing]
  The Police looked for evidence.        [Police looked evidence _ _]
  The Police found that Tony bomb(ed) City-Hall.
                               [Police found City-Hall bomb(ed) Tony]
  The Police was after Tony.                [Police was _ after Tony]
[$being-after Police Tony bomb(ed) arrest(ed) New-York City-Hall _]
  The Police was after Tony.                [Police was _ after Tony]
  The Police thought that Tony bomb(ed) City-Hall.
                             [Police thought City-Hall bomb(ed) Tony]
  The Police wanted to arrest(ed) Tony.
                                    [Police wanted _ arrest(ed) Tony]
  The Police found that Tony was in New-York.
                                     [Police found New-York was Tony]
  The Police planned to arrest(ed) Tony in New-York.
                            [Police planned New-York arrest(ed) Tony]
[$driving Tony _ scared Chicago Chicago recklessly _]
  Tony wanted to go to Chicago.            [Tony wanted Chicago go _]
  Tony entered his car.                      [Tony entered _ his car]
  Tony drove to Chicago.                     [Tony drove Chicago _ _]
  Tony was scared.                              [Tony was _ _ scared]
  Tony drove recklessly.                [Tony drove _ _ recklessly]
[$job Fred Police worked liked New-York good cop]
  Fred was a cop.                                  [Fred was _ _ cop]
  Fred worked for the Police.                [Fred worked Police _ _]
  Fred worked in New-York.                 [Fred worked New-York _ _]
  Fred liked his job.                          [Fred liked _ his job]
  Fred was a good cop.                          [Fred was _ good cop]
```

```
[$pulled-over Tony Fred arrest(ed) _ bombing _ _]
  Tony was pulled-over by Fred.         [Tony was Fred _ pulled-over]
  Fred asked Tony for his license.      [Fred asked license his Tony]
  Tony gave his license to Fred.         [Tony gave Fred his license]
  Fred checked the license.                [Fred checked _ _ license]
  Fred arrest(ed) Tony for bombing.  [Fred arrest(ed) bombing _ Tony]
[$talking Tony Fred hated liked hand-shake Vito long]
  Tony talked to Fred about Vito.       [Tony talked Fred about Vito]
  Tony hated Vito.                                [Tony hated _ _ Vito]
  Tony talked to Fred a long time.       [Tony talked Fred long time]
  Tony liked talking to Fred.             [Tony liked Fred _ talking]
  Tony gave a hand-shake good-bye to Fred.
                                 [Tony gave Fred hand-shake good-bye]
[$being-after Police Vito bomb(ed) arrest(ed) New-York City-Hall _]
  The Police was after Vito.               [Police was _ after Vito]
  The Police thought that Vito bomb(ed) City-Hall.
                             [Police thought City-Hall bomb(ed) Vito]
  The Police wanted to arrest(ed) Vito.
                                    [Police wanted _ arrest(ed) Vito]
  The Police found that Vito was in New-York.
                                     [Police found New-York was Vito]
  The Police planned to arrest(ed) Vito in New-York.
                            [Police planned New-York arrest(ed) Vito]
story-end


Story #24

Emotion: neutral      Context: Gangster
[$relation Fred Bob liked trusted _ Fred boss]
  Bob was the boss of Fred.                     [Bob was Fred _ boss]
  Fred liked Bob.                                 [Fred liked _ _ Bob]
  Fred trusted Bob.                             [Fred trusted _ _ Bob]
[$plan Bob Fred praise(d) City-Hall Meeting Fred invitation]
  Bob planned a Meeting at City-Hall.
                                    [Bob planned City-Hall _ meeting]
  Bob wanted to praise(d) Fred.         [Bob wanted _ praise(d) Fred]
  Bob sent invitation to Fred.           [Bob sent Fred _ invitation]
  Fred accepted the invitation.        [Fred accepted _ _ invitation]
```

```
[$driving Fred _ on-time City-Hall City-Hall carefully _]
  Fred wanted to go to City-Hall.        [Fred wanted City-Hall go _]
  Fred entered his car.                      [Fred entered _ his car]
  Fred drove to City-Hall.                 [Fred drove City-Hall _ _]
  Fred was on-time.                            [Fred was _ _ on-time]
  Fred drove carefully.                    [Fred drove _ _ carefully]
[$occasion Fred Bob speech praise(d) City-Hall Fred meeting]
  Fred entered City-Hall for meeting.[Fred entered meeting _ City-H.]
  Bob praise(d) Fred.                        [Bob praise(d) _ _ Fred]
  The meeting was a success.                [meeting was _ _ success]
  Fred gave a speech.                         [Fred gave _ _  speech]
  Fred drank champagne.                    [Fred drank _ _ champagne]
[$drunk Fred _ _ _ champagne good a-little]
  Fred ordered more champagne.       [Fred ordered _ more champagne]
  Fred got a-little drunk.                [Fred got _ a-little drunk]
  Fred had a good time.                        [Fred had _ good time]
story-end

Story #25

Emotion: very negative      Context: Gangster
[$job Vito Mafia head liked New-York famous gangster]
  Vito was a gangster.                        [Vito was _ _ gangster]
  Vito was the head of the Mafia.             [Vito was Mafia _ head]
  Vito worked in New-York.                 [Vito worked New-York _ _]
  Vito liked his job.                           [Vito liked _ his job]
  Vito was a famous gangster.          [Vito was _ famous gangster]
[$driving Vito _ scared airport LA recklessly _]
  Vito wanted to go to LA.                      [Vito wanted LA go _]
  Vito entered his car.                      [Vito entered _ his car]
  Vito drove to the airport.                [Vito drove airport _ _]
  Vito was scared.                              [Vito was _ _ scared]
  Vito drove recklessly.                  [Vito drove _ _ recklessly]
[$pulled-over Vito cop arrest(ed) _ bombing _ _]
  Vito was pulled-over by a cop.         [Vito was cop _ pulled-over]
  The cop asked Vito for his license.    [cop asked license his Vito]
  Vito gave his license to The cop.       [Vito gave cop his license]
  The cop checked the license.              [cop checked _ _ license]
  The cop arrest(ed) Vito for bombing.[cop arrest(ed) bombing _ Vito]
```

```
[$trial Vito _ walked cleared free bombing good]
  Vito was accused of bombing.           [Vito was bombing _ accused]
  Vito was brought before the court.       [Vito was court _ brought]
  Vito had a good lawyer.                    [Vito had _ good lawyer]
  The court cleared Vito of bombing.   [court cleared bombing _ Vito]
  Vito walked free.                            [Vito walked _ free _]
story-end


Story #26

Emotion: very negative      Context: Gangster
[$being-after Mafia Tony talked kill(ed) New-York Police _]
  The Mafia was after Tony.                  [Mafia was _ after Tony]
  The Mafia thought that Tony talked to Police.
                                   [Mafia thought Police talked Tony]
  The Mafia wanted to kill(ed) Tony.   [Mafia wanted _ kill(ed) Tony]
  The Mafia found that Tony was in New-York.
                                      [Mafia found New-York was Tony]
  The Mafia planned to kill(ed) Tony in New-York.
                               [Mafia planned New-York kill(ed) Tony]
[$driving Vince _ on-time airport New-York carefully _]
  Vince wanted to go to New-York.        [Vince wanted New-York go _]
  Vince entered his car.                    [Vince entered _ his car]
  Vince drove to the airport.               [Vince drove airport _ _]
  Vince was on-time.                          [Vince was _ _ on-time]
  Vince drove carefully.                  [Vince drove _ _ carefully]
[$flight Vince no-one _ LA New-York short late]
  Vince entered the LA airport.          [Vince entered _ LA airport]
  Vince walked to the counter.             [Vince walked counter _ _]
  Vince checked his bag.                    [Vince checked _ his bag]
  Vince walked to his gate.                 [Vince walked gate his _]
  Vince entered the plane to New-York.
                                     [Vince entered New-York _ plane]
  The plane was late.                             [plane was _ _ late]
  Vince met no-one in the plane.           [Vince met plane _ no-one]
[$plan Vito Tony kill(ed) New-York murder Vince order]
  Vito planned a murder in New-York. [Vito planned New-York _ murder]
  Vito wanted to kill(ed) Tony.         [Vito wanted _ kill(ed) Tony]
```

```
  Vito gave order to Vince.                [Vito gave Vince _ order]
  Vince accepted the order.               [Vince accepted _ _ order]
[$driving Vince _ on-time Tony Tony carefully _]
  Vince wanted to go to Tony.               [Vince wanted Tony go _]
  Vince entered his car.                   [Vince entered _ his car]
  Vince drove to Tony.                        [Vince drove Tony _ _]
  Vince was on-time.                         [Vince was _ _ on-time]
  Vince drove carefully.                 [Vince drove _ _ carefully]
[$occasion Vince Vince phone-call kill(ed) Starbucks Tony murder]
  Vince entered Starbucks for murder.
                                    [Vince entered murder _ Starbucks]
  Vince kill(ed) Tony.                     [Vince kill(ed) _ _ Tony]
  The murder was a success.                 [murder was _ _ success]
  Vince made a phone-call.               [Vince made _ _  phone-call]
  Vince smoked a cigarette.              [Vince smoked _ _ cigarette]
[$investigation nothing Police _ Starbucks _ murder _]
  The Police investigated the murder at Starbucks.
                              [Police investigated Starbucks _ murder]
  The Police looked for evidence.        [Police looked evidence _ _]
  The Police found nothing.               [Police found _ _ nothing]
  The Police was after no-one.           [Police was _ after no-one]
story-end
```

```
Story #27

Emotion: very negative      Context: Gangster
[$drink Vince Vito  met Starbucks table coffee _]
  Vince went to Starbucks.                [Vince went Starbucks _ _]
  Vince sat at a table.                        [Vince sat table _ _]
  Vince ordered coffee.                   [Vince ordered _ _ coffee]
  Vince drank the coffee.                   [Vince drank _ _ coffee]
  Vince met Vito at Starbucks.          [Vince met Starbucks _ Vito]
[$relation Vince Vito liked feared _ Vince Boss]
  Vito was the Boss of Vince.               [Vito was Vince _ Boss]
  Vince liked Vito.                           [Vince liked _ _ Vito]
  Vince feared Vito.                         [Vince feared _ _ Vito]
```

```
[$plan Vito Bob kill(ed) New-York murder Vince order]
  Vito planned a murder in New-York. [Vito planned New-York _ murder]
  Vito wanted to kill(ed) Bob.           [Vito wanted _ kill(ed) Bob]
  Vito gave order to Vince.                 [Vito gave Vince _ order]
  Vince accepted the order.               [Vince accepted _ _ order]
[$talking Vito Vince liked liked hand-shake murder short]
  Vito talked to Vince about murder. [Vito talked Vince about murder]
  Vito liked murder.                           [Vito liked _ _ murder]
  Vito talked to Vince a short time.   [Vito talked Vince short time]
  Vito liked talking to Vince.           [Vito liked Vince _ talking]
  Vito gave a hand-shake good-bye to Vince.
                             [Vito gave Vince hand-shake good-bye]
[$driving Vince _ scared City-Hall City-Hall carefully _]
  Vince wanted to go to City-Hall.      [Vince wanted City-Hall go _]
  Vince entered his car.                    [Vince entered _ his car]
  Vince drove to City-Hall.               [Vince drove City-Hall _ _]
  Vince was scared.                             [Vince was _ _ scared]
  Vince drove carefully.                  [Vince drove _ _ carefully]
[$occasion Vince Vince phone-call kill(ed) City-Hall Bob murder]
  Vince entered City-Hall for murder.
                                [Vince entered murder _ City-Hall]
  Vince kill(ed) Bob.                         [Vince kill(ed) _ _ Bob]
  The murder was a success.                 [murder was _ _ success]
  Vince made a phone-call.               [Vince made _ _  phone-call]
  Vince smoked a cigarette.              [Vince smoked _ _ cigarette]
story-end
```

Story #28

```
Emotion: negative      Context: Gangster
[$investigation Vince Police kill(ed) City-Hall Bob murder _]
  The Police investigated the murder at City-Hall.
                            [Police investigated City-Hall _ murder]
  The Police looked for evidence.        [Police looked evidence _ _]
  The Police found that Vince kill(ed) Bob.
                                 [Police found Bob kill(ed) Vince]
  The Police was after Vince.              [Police was _ after Vince]
```

```
[$being-after Police Vince kill(ed) arrest(ed) New-York Bob _]
  The Police was after Vince.              [Police was _ after Vince]
  The Police thought that Vince kill(ed) Bob.
                                  [Police thought Bob kill(ed) Vince]
  The Police wanted to arrest(ed) Vince.
                                   [Police wanted _ arrest(ed) Vince]
  The Police found that Vince was in New-York.
                                    [Police found New-York was Vince]
  The Police planned to arrest(ed) Vince in New-York.
                           [Police planned New-York arrest(ed) Vince]
[$driving Vince _ scared airport LA recklessly _]
  Vince wanted to go to LA.                    [Vince wanted LA go _]
  Vince entered his car.                    [Vince entered _ his car]
  Vince drove to the airport.               [Vince drove airport _ _]
  Vince was scared.                            [Vince was _ _ scared]
  Vince drove recklessly.                [Vince drove _ _ recklessly]
[$pulled-over Vince cop arrest(ed) _ murder _ _]
  Vince was pulled-over by a cop.       [Vince was cop _ pulled-over]
  The cop asked Vince for his license.  [cop asked license his Vince]
  Vince gave his license to the cop.     [Vince gave cop his license]
  The cop checked the license.              [cop checked _ _ license]
  The cop arrest(ed) Vince for murder.[cop arrest(ed) murder _ Vince]
[$trial  Vince _ went convicted jail murder good]
  Vince was accused of murder.           [Vince was murder _ accused]
  Vince was brought before the court.     [Vince was court _ brought]
  Vince had a good lawyer.                  [Vince had _ good lawyer]
  The court convicted Vince of murder.
                                    [court convicted murder _ Vince]
  Vince went to jail.                           [Vince went jail _ _]
story-end
```

37

Bibliography

Abitz, M., Nielsen, R., Jones, E., Laursen, H., Graem, N., and Pakkenberg, B. (2007). Excess of neurons in the human newborn mediodorsal thalamus compared with that of the adult. *Cerebral Cortex*, 17(11):2573–8.

Abolafia, M. (2010). Narrative construction as sensemaking. *Organization Studies*, 31(3):349–367.

Adeli, H., Ghosh-Dastidar, S., and Dadmehr, N. (2005). Alzheimer's disease and models of computation: Imaging, classification, and neural models. *Journal of Alzheimer's Disease*, 7(3):187–199.

Aleman, A., Hijman, R., de Haan, E., and Kahn, R. (1999). Memory impairment in schizophrenia: A meta-analysis. *American Journal of Psychiatry*, 156:1358–1366.

Aloia, M. S., Gourovitch, M. L., Missar, D., Pickar, D., Weinberger, D. R., and Goldberg, T. E. (1998). Cognitive substrates of thought disorder, II: specifying a candidate cognitive mechanism. *American Journal of Psychiatry*, 155:1677–84.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders (DSM IV-TR)*. American Psychiatric Association, fourth edition.

Andreasen, N. (1979). Thought, language, and communication disorders. II. diagnostic significance. *Archives of General Psychiatry*, 26:1325–30.

Andreasen, N. (1984). *The Scale for the Assessment of Positive Symptoms (SAPS)*. University of Iowa, Iowa City.

Andreasen, N. (1987). *Comprehensive assessment of symptoms and history (CASH)*. University of Iowa, Iowa City.

Appelbaum, P., Robbins, P., and Roth, L. (1999). Dimensional approach to delusions: Comparison across types and diagnoses. *American Journal of Psychiatry*, 156:1938–1943.

Arnsten, A. (2007). Catecholamine and second messenger influences on prefrontal cortical networks of "representational knowledge": A rational bridge between genetics and the symptoms of mental illness. *Cerebral Cortex*, 17 Suppl 1:i6–15.

Arnsten, A., and Goldman-Rakic, P. (1998). Noise stress impairs prefrontal cortical cognitive function in monkeys: evidence for a hyperdopaminergic mechanism. *Archives of General Psychiatry*, 55(4):362–8.

Assadi, S., Yücel, M., and Pantelis, C. (2009). Dopamine dysregulation in the brain network of decision-making: can this explain the psychopathology of schizopohrenia? In Kasper, S., and Papadimitriou, G., editors, *Schizophrenia. Biophysical Approaches and Current Challenges*, Medical Psychiatry Series, chapter Chapter 12, 131–151. New York: Informa Healthcare, second edition.

Assaf, M., Rivkin, P., CKuzu, C., Calhoun, V., Kraut, M., Groth, K., Yassa, M., Hart, J., and Pearlson, G. (2006). Abnormal object recall and anterior cingulate overactivation correlate with formal thought disorder in schizophrenia. *Biological Psychiatry*, 59:452–9.

Bentall, R. P., Rowse, G., Shryane, N., Kinderman, P., Howard, R., Blackwood, N., Moore, R., and Corcoran, R. (2009). The cognitive and affective structure of paranoid delusions: a transdiagnostic investigation of patients with schizophrenia spectrum disorders and depression. *Archives of General Psychiatryry*, 66:236–47.

Berridge, K., and Robinson, T. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Review*, 28:309–69.

Bertsekas, D. P., and Tsitsiklis, J. N. (1996). *Neuro-Dynamic Programming*. Belmont, MA: Athena Scientific. ISBN 1-886529-10-8.

Bilder, R., Goldman, R., and Robinson, D. e. a. (2000). Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. *American Journal of Psychiatry*, 157(4):549–59.

Bleuler, E. (1911). Dementia praecox oder die Gruppe der Schizofrenien. In Aschaffenburg, G., editor, *Handbuch der Psychiatrie*. Leipzig: Breithep und Hartel.

Bogerts, B., Steiner, J., and Bernstein, H. (2009). Brain abnormalities in schizophrenia. In Kasper, S., and Papadimitriou, G., editors, *Schizophrenia. Biophysical Approaches and Current Challenges*, Medical Psychiatry Series, chapter Chapter 9, 87–104. New York: Informa Healthcare, second edition.

Bower, G., and Morrow, D. (1990). Mental models in narrative comprehension. *Science*, 247:44–48.

Braver, T., Barch, D., and J, C. (1999). Cognition and control in schizophrenia: A computational model of dopamine and prefrontal function. *Biological Psychiatry*, 46:312–28.

Brewer, W. (2005). Memory impairments identified in people at ultra high-risk for psychosis who later develop first episode psychosis. *American Journal of Psychiatry*, 162:71–78.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, 18:1–21.

Buchanan, R., Vladar, K., Barta, P., and Pearlson, G. (1998). Structural evaluation of the prefrontal cortex in schizophrenia. *American Journal of Psychiatry*, 155:1049–55.

Caramazza, A., Hillis, A., Leek, E., and Miozzo, M. (1994). The organization of lexical knowledge in the brain: Evidence from category- and modality-specific deficits. In Hirschfeld, L., and Gelman, S., editors, *Mapping the Mind*, 68–84. Cambridge University Press.

Carter, J., and Neufeld, R. (2007). Cognitive processing of facial affect: Connectionist model of deviations in schizophrenia. *Journal of Abnormal Psychology*, 116:90–305.

Centorrino F, e. a. (2002). Inpatient antipsychotic drug use in 1998, 1993, and 1989. *American Journal of Psychiatry*, 159:1932–5.

Cerf, B. (1993). The gift. In Carfield, J., and Hansen, M., editors, *Chicken Soup for the Soul*, p23. Deerfield Beach, FL: Health Communications, Inc.

Chapin, K., Vann, L., Lycaki, H., Josef, N., and Meyendorff, E. (1989). Investigation of the semantic network in schizophrenia using the semantic priming paradigm. *Schizophrenia Research*, 2:355–60.

Chechik, G., Meilijson, I., and Ruppin, E. (1998). Neuronal regulation: A mechanism for synaptic pruning during brain maturation. *Neural Computation*, 11:11–8.

Christiansen, M., Chater, N., and Seidenberg, M. E. (1999). Connectionist models of human language processing: Progress and prospects. *Special issue of Cognitive Science*, 24(4):415–63.

Cohen, J., Braver, T., and O'Reilly, R. (1996). A computational approach to prefrontal cortex, congnitive control, and schizophrenia: Recent developments and current challenges. *Philosophical Transactions of the Royal Society of London Series B (Biological Sciences)*, 351:1515–27.

Cohen, J., and Servan-Schreiber, D. (1992). Context, cortex and dopamine: A connectionist approach to behaviour and biology in schizophrenia. *Psychology Review*, 99:45–77.

Colby, K. (1973). Simulations of belief systems. In Schank, R., and Colby, K., editors, *Computer Models of Thought and Language*. W. H. Freeman and Company.

Collobert, R., and Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25 th International Conference on Machine Learning*. Helsinki, Finland.

Corlett, P. e. a. (2006). Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. *Archives of General Psychiatryry*, 63:611–21.

Covington, M., He, C., Brown, C., Naci, L., McClain, J., Fjordbak, B., Semple, J., and Brown, J. (2005). Schizophrenia and the structure of language: The linguist's view. *Schizophrenia Research*, 77:85–98.

Crockett, D., Tallman, K., Hurwitz, T., and Kozak, J. (1988). Neuropsychological performance in psychiatric patients with or without documented brain dysfunction. *International Journal of Neuroscience*, 41:71–9.

Crow, T. (1990). Temporal lobe asymmetries as the key to the etiology of schizophrenia. *Schizophrenia Bulletin*, 16:433–43.

Cullingford, R. (1978). Script application: Computer understanding of newspaper stories. Doctoral Dissertation, Department of Computer Science, Yale University.

Curran, C., Byrappa, N., and Mcbride, A. (2004). Stimulant psychosis: systematic review. *British Journal of Psychiatry*, 185:196–204.

Davis, J. (1974). Dose equivalence of the antipsychotic drugs. *Journal of Psychiatric Research*, 11:65–69.

Dazzi, L., Seu, E., Cherchi, G., and Biggio, G. (2004). Inhibition of stress-induced dopamine output in the rat prefrontal cortex by chronic treatment with olanzapine. *Biological Psychiatry*, 55(5):477–83.

Den Uyl, M., and van Oostendorp, H. (1980). The use of scripts in text comprehension. *Poetics*, 9:275–94.

Dennett, D. C. (1992). The self as a center of narrative gravity. In Kessel, F., Cole, P., and Johnson, D., editors, *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Erlbaum.

D'Esposito, M., Postle, B., and Rypma, B. (2000). Prefrontal cortical contributions to working memory: evidence from event-related fMRI studies. *Experimental Brain Research*, 133:3–11.

Docherty, N., Hall, M., and Gordinier, S. (1998). Affective reactivity of speech in schizophrenic patients and their nonschizophrenic relatives. *Journal of Abnormal Psychology*, 107:461–467.

Egan, M., and Weinberger, D. (1997). Neurobiology of schizophrenia. *Opinions in Neurobiology*, 7:701–707.

Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14:179–211.

Euston, D., Tatsuno, M., and McNaughton, B. (2007). Fast-forward playback of recent memory sequences in prefrontal cortex during sleep. *Science*, 318:5853.

Farah, M., and Wallace, M. (1992). Semantically bounded anomia: Implications for the neural implementation of naming. *Neuropsychologia*, 30:609–21.

Farkas, I., and Crocker, M. (2008). Syntactic systematicity in sentence processing with a recurrent self-organizing network. *Neurocomputing*, 71:1172–1179.

Feinberg, I. (1982/1983). Schizophrenia caused by a fault in programmed synaptic elimination during adolescence? *Journal of Psychiatric Research*, 17:319–34.

Fidelman, P., Miikkulainen, R., and Hoffman, R. E. (2005). A subsymbolic model of complex story understanding. In *Proceedings of CogSci'05*.

Finkel, L. (2000). Neuroengineering models of brain disease. *Annu Rev Biomed Eng*, 2:577–606.

First, M., Spitzer, R., Gibbon, M., and Williams, J. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders - Non-patient Edition*. New York Psychiatric Institute, New York City, NY.

Flanagan, O. (1992). *Consciousness Reconsidered*. MIT Press.

Fox Tree, J. (2000). Coordinating spontaneous talk. In Wheeldon, L., editor, *Aspects of language production*, 375–406. New York, NY: Psychology Press.

Gallup, G., and Newport, F. (1991). Belief in paranormal phenomena among adult Americans. *Skeptical Inquirer*, 15:137–146.

Glantz, L., Gilmore, J., Lieberman, J., and Jarskog, L. (2006). Apoptotic mechanisms and the synaptic pathology of schizophrenia. *Schizophrenia Research*, 81:47–63.

Glantz, L., and Lewis, D. (2000). Decreased dendritic spine density on prefrontal cortical pyramidal neurons in schizophrenia. *Archives of General Psychiatry*, 57:65–73.

Glenthoj, B., Christiansen, L., Rasmussen, H., and Oranje, B. (2009). Biochemical slterations in schizophrenia. In Kasper, S., and Papadimitriou, G., editors, *Schizophrenia. Biophysical Approaches and Current Challenges*, Medical Psychiatry Series, chapter Chapter 11, 114–130. New York: Informa Healthcare, second edition.

Goldberg, T., Aloia, M., Gourovitch, M., Missar, D., Pickar, D., and Weinberger, D. (1998). Cognitive substrates of thought disorder, I: the semantic system. *American Journal of Psychiatry*, 155:1671–6.

Goldman-Rakic, P., Castner, S., Svensson, T., Siever, L., and Williams, G. (2004). Targeting the dopamine D1 receptor in schizophrenia: insights for cognitive dysfunction. *Psychopharmacology*, 174(1):3–16.

Graesser, A., Woll, S., Kowalski, D., and Smith, D. (1980). Memory for typical and atypical actions in scripted activities. *Journal of Experimantal Psychol-Hum L*, 6:503–15.

Grasemann, U., Kiran, S., Sandberg, C., and Miikkulainen, R. (2011). Simulating bilingual aphasia: a novel computational model. In preparation.

Grasemann, U., Miikkulainen, R., and Hoffman, R. (2007). A subsymbolic model of language pathology in schizophrenia. In *Proceedings of CogSci'07*. Nashville, Tennessee, USA.

Grasemann, U., Miikkulainen, R., and Hoffman, R. (2009). Hyperlearning: A connectionist model of psychosis in schizophrenia. In *Proceedings of CogSci'09*. Amsterdam, The Netherlands.

Grossberg, S. (1999). How hallucinations may arise from brain mechanisms of learning, attention, and volition. *Journal of the International Neuropsychological Society*, 6:579–588.

Grossberg, S. (2000). The imbalanced brain: from normal behavior to schizophrenia. *Biological Psychiatry*, 48:81–98.

Grossberg, S., and Pepe, J. (1970). Schizophrenia: possible dependence of associational span, bowing and primacy versus recency on spiking threshold. *Behav Sci*, 15:359–62.

Guo, N., Hwang, D., Lo, E., Huang, Y., Laruelle, M., and Abi-Dargham, A. (2003). Dopamine depletion and in vivo binding of PET D1 receptor radioligands: implications for imaging studies in schizophrenia. *Neuropsychopharmacology*, 28(9):1703–11.

Gur, R., Calkins, M., and Gur, R. e. a. (2007). The consortium on the genetics of schizophrenia: neurocognitive endophenotypes. *Schizophrenia Bulletin*, 33:49–68.

Hampton, A., and O'Doherty, J. (2007). Decoding the neural substrates of rewaed-related decision making with functional MRI. *Proc Natl Acad Sci USA*, 104(4):1377–82.

Harm, M., and Seidenberg, M. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, 106(3):491–528.

Harrow, M., Herbener, E., Shanklin, A., Jobe, T., Rattenbury, F., and Kaplan, K. (2004). Followup of psychotic outpatients: dimensions of delusions and work functioning in schizophrenia. *Schizophrenia Bulletin*, 30:147–61.

Heckers, S., Rauch, S., Goff, D., Savage, C., Schachter, D., Fischman, A., and Alpert, N. (1998). Impaired recruitment of the hippocampus during conscious recollection in schizophrenia. *Nature Neuroscience*, 1:318–323.

Hinton, G., editor (1991). *Connectionist Symbol Processing*. Cambridge, MA: MIT Press.

Hoffman, R. (1987). Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Archives of General Psychiatry*, 44:178–188.

Hoffman, R., Grasemann, U., Gueorguieva, R., Quinlan, D., Lane, D., and Miikkulainen, R. (2010). Using computational patients to evaluate illness mechanisms in schizophrenia. In review.

Hoffman, R., and McGlashan, T. (1997). Synaptic elimination, neurodevelopment and the mechanism of hallucinated ‘voices’ in schizophrenia. *American Journal of Psychiatry*, 154:1683–1689.

Hoffman, R., and McGlashan, T. (2006). Using a speech perception neural network computer simulation to contrast neuroanatomic versus neuromodulatory models of auditory hallucinations. *Pharmacopsychiatry*, 39 (suppl 1):554–64.

Hoffman, R., and Sledge, W. (1988). An analysis of grammatical deviance occurring in the spontaneous speech of schizophrenics. *Journal of Neurolinguistics*, 3:89–101.

Hoffman, R., Stopek, S., and Andreasen, N. (1986). A comparative study of manic versus schizophrenic speech disorganization. *Archives of General Psychiatry*, 43:831–8.

Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci*, 79:2554–58.

Hunter, J. (2007). Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, 9(3):90–95.

Huttenlocher, P. (1979). Synaptic density in the human frontal cortex – developmental changes and effects of aging. *Brain Research*, 163:195–205.

Huttenlocher, P. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, 387:167–78.

Huys, Q. (2007). *Reinforcers and control: Towards a computational aetiology of depression*. Ph.D. thesis, Gatsby Unit, University College London.

Jhala, A., and van Velsen, M. (2009). Challenges in development and design of interactive narrative authoring systems. In *Proceedings of the AAAI Spring Symposium on Intelligent Narrative Technologies II*. Palo Alto, California.

Kane, J. (1997). Update on treatment strategies. *Int Rev Psych*, 9:419–27.

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, 160(1):13–23.

Kapur, S., and Mamo, D. (2003). Half a century of antipsychotics and still a central role dopamine D2 receptors. *Prog Neuropsychopharmacol Biol Psychiat*, 27(7):1081–90.

Kapur, S., Mizrahi, R., and Li, M. (2005). From dopamine to salience to psychosis – linking biology, pharmacology and phenomenology of psychosis. *Schizophrenia Research*, 79(1):919–28.

Kapur, S., and Seeman, P. (2001). Does fast dissociation from the dopamine d(2) receptor explain the action of atypical antipsychotics? *American Journal of Psychiatry*, 158:360–9.

Karlsgodt, K. e. a. (2008). Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biological Psychiatry*, 63:512–8.

Kellendonk, C., Simpson, E., Polan, H., Malleret, G., Vronskaya, S., Winiger, V., Moore, H., and Kandel, E. (2006). Transient and selective overexpression of dopamine D2 receptors in the striatum causes persistent abnormalities in prefrontal cortex functioning. *Neuron*, 49(4):603–15.

Kemp, R., Chua, S., McKenna, P., and David, A. (1997). Reasoning and delusions. *British Journal of Psychiatry*, 170:398–405.

Kennerley, S., Walton, M., Behrens, T., Buckley, M., and Rushworth, M. (2006). Optimal decision making and the anterior congulate cortex. *Nature Neuroscience*, 9(7):940–7.

Kerns, J., and Berenbaum, H. (2003). The relationship between formal thought disorder and executive functioning component processes. *Journal of Abnormal Psychology*, 112:339–52.

Kim, J., Kwon, J., Park, H., Youn, T., Kang, D., Kim, M., Lee, D., and Lee, M. (2003). Functional disconnection between the prefrontal and parietal cortices during working memory processing in schizophrenia: a [15(O)]H2O PET study. *American Journal of Psychiatry*, 160:919–23.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69.

Kolodner, J. (1993). *Case-Based Reasoning*. San Mateo, CA: Morgan Kaufmann Publishers.

Kornetsky, C., and Eliasson, M. (1969). Reticular stimulation and chlorpromazine: An animal model for schizophrenic overarousal. *Science*, 165(3899):1273–4.

Kraepelin, E. (1896). *Psychiatrie. Ein Lehrbuch für Studirende und Aerzte [Psychiatry. A textbook for students and physicians]*. Leipzig: Verlag von Johann Ambrosius Barth, fünfte, vollständig umgearbeitete auflage. [5th revised] edition.

Kranz, H. (1967). Wahn und zeitgeist. *Studium Generale*, 20:606–611.

Kraus, M., Keefe, R., and Krishnan, R. (2009). Memory-prediction errors and their consequences in schizophrenia. *Neuropsychological Review*, 19:336–52.

Kuperberg, G. (2010). Language in schizophrenia part 1: An introduction. *Language and Linguistics Compass*, 4(8):576–589.

Kuperberg, G. R., Deckersbach, T., Holt, D. J., Goff, D., and West, W. C. (2007). Increased temporal and prefrontal activity in response to semantic associations in schizophrenia. *Archives of General Psychiatryry*, 64:138–151.

Kwapil, T., Hegley, D., Chapman, L., and Chapman, J. (1990). Facilitation of word recognition by semantic priming in schizophrenia. *Journal of Abnormal Psychology*, 99:215–21.

Laruelle, M. (2000). The role of endogenous sensitization in the pathophysiology of schizophrenia: implications from recent brain imaging studies. *Brain Research - Brain Research Reviews*, 31(2–3):371–84.

Leeson, V., Simpson, A., McKenna, P., and Laws, K. (2005). Executive inhibition and semantic association in schizophrenia. *Schizophrenia Research*, 74:61–7.

Lewis, D., and Gonzales-Burgos, G. (2006). Pathophysiologically based treatment interventions in schizophrenia. *Nature Medicine*, 12(9):1016–22.

LF, J., Glantz, L., Gilmore, J., and Lieberman, J. (2005). Apoptotic mechanisms in the pathophysiology of schizophrenia. *Progress in Neuropsychopharmacology & Biological Psychiatry*, 29:846–58.

Li, P., Farkas, I., and MacWhinney, B. (2004). Early lexical acquisition in a self-organizing neural network. *Neural Networks*, 17:1345–1362.

Lidow, M., Elsworth, J., and Goldman-Rakic, P. (1997). Down-regulation of the D1 and D5 dopamine receptors in the primate prefrontal cortex by chronic treatment with antipsychotic drugs. *Journal of Pharmacology and Experimental Therapeutics*, 281(1):597–603.

Lieberman, J., Kane, J., and Alvir, J. (1987). Provocative tests with psychostimulant drugs in schizophrenia. *Psychopharmacology*, 91(4):415–33.

Loh, M., Rolls, E., and Deco, G. (2007). A dynamical systems hypothesis of schizophrenia. *PLOS Comput Biology*, 3:e228.

Maher, B. (1974). Delusional thinking and perceptual disorder. *Journal of Individual Psychology*, 30:98–113.

Maher, B. (2002). Delusions. In Adams, H., and Sutker, B., editors, *Comprehensive hand-*

book of psychopathology, chapter Chapter 12, 309–340. New York, NY: Kluwer Academic/Plenum Publishers, third edition.

Maher, B., and Deldin, P. (2002). Schizophrenia: Biopsychological aspects. In Adams, H., and Sutker, B., editors, *Comprehensive handbook of psychopathology*, chapter Chapter 13, 341–370. New York, NY: Kluwer Academic/Plenum Publishers, third edition.

Maher, B., Manschreck, T., Hoover, T., and Weisstein, C. (1987). Thought disorder and measured features of language production. In Harvey, P., and Walker, E., editors, *Positive and Negative Symptoms in Psychosis*, 195–215. Hillsdale, NJ: Erlbaum Associates.

Manschreck, B., Maher, B., Milavetz, J., Ames, D., Weisstein, C., and Schneyer, M. (1988). Semantic priming in thought-disordered schizophrenic patients. *Schizophrenia Research*, 1:61–66.

Mayberry, M., and Miikkulainen, R. (2005). Broad-coverage parsing with neural networks. *Neural Processing Letters*, 21:121–132.

McClelland, J., McNaughton, B., and O'Reilly, R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102:419–457.

McClelland, J., Rumelhart, D., and the PDP Research Group, editors (1987). *Parallel Distributed Processing, Vol. 2: Psychological and Biological Models*. Cambridge, MA: MIT Press.

McGaugh, J. (2000). Memory: a century of consolidation. *Science*, 287:248–251.

McGuire, P. e. a. (1998). Pathophysiology of 'positive' thought disorder in schizophrenia. *Br J Psychiatry*, 173:231–5.

Meyer-Lindenberg, A., Poline, J.-B., Kohn, P. D., Holt, J. L., Egan, M. F., Weinberger, D. R., and Berman, K. F. (2001). Evidence for abnormal cortical functional connectivity during working memory in schizophrenia. *American Journal of Psychiatry*, 158:1809–17.

Miikkulainen, R. (1990). A PDP architecture for processing sentences with relative clauses. In *In Proceedings of the 13th International Conference on Computational Linguistics*, 3–201.

Miikkulainen, R. (1993). *Subsymbolic Natural Language Processing: An Integrated Model of Scripts, Lexicon, and Memory*. Cambridge, MA: MIT Press.

Miikkulainen, R. (1996). Subsymbolic case-role analysis of sentences with embedded clauses. *Cog Sci*, 20:47–73.

Miikkulainen, R. (1997). Dyslexic and category-specific aphasic impairments in a self-organizing feature map model of the lexicon. *Brain and Language*, 59:334–66.

Miikkulainen, R., Bednar, J., Choe, Y., and Sirosh, J. (2005). *Computational Maps in the Visual Cortex*. New York, NY: Springer.

Miikkulainen, R., and Dyer, M. (1991). Natural language processing with modular PDP networks and distributed lexicon. *Cognitive Science*, 15:343–399.

Miikkulainen, R., and Kiran, S. (2009). Modeling the bilingual lexicon of an individual subject. In Principe, J., and Miikkulainen, R., editors, *Advances in Self-Organizing Maps*, Lecture Notes in Computer Science, 191–199. Springer Berlin / Heidelberg.

Monchi, O., Taylor, T., and Dagher, A. (2000). A neural model of working memory processes in normal subjects, parkinson's disease and schizophrenia for fMRI design and predictions. *Neural Networks*, 13:953–73.

Moritz, S., Woodward, T., Kuppers, D., Lausen, A., and Schickel, M. (2003). Increased automatic spreading of activation in thought-disordered schizophrenic patients. *Schizophr Res*, 59:181–6.

Neely, J. (1977). Semantic priming and retrieval from lexical memory: roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology*, 106:226 254.

Neely, J. (1991). Semantic priming effects in visual word recognition: a selective review of current findings and theories. In Besner, D., and Humphreys, G., editors, *Basic Progresses in Reading and Visual Word Recognition*, 264–333. Hillsdale, NJ: Erlbaum.

Noll, R. (2006). *Encyclopedia of Schizophrenia and Other Psychotic Disorders*. New York, NY: Facts on File, third edition.

Nuechterlein, K., and Dawson, M. (1984). Information processing and attentional functioning in the developmental course of schizophrenic disorders. *Schizophrenia Bulletin*, 10:160–203.

Ober, B., Vinogradov, S., and Shenaut, G. (1995). Semantic priming of category relations in schizophrenia. *Neuropsychology*, 9:220 228.

O'Reilly, R., and Frank, M. (2006). Making working memory work: A computational model of learning in the frontal cortex and basal ganglia. *Neural Computation*, 18:283–328.

Pakkenberg, B. (1993). Total nerve cell number in neocortex estimated using optical disectors. *Biological Psychiatry*, 34:768–72.

Pearlson, G. (2000). Neurobiology of schizophrenia. *Ann Neurol*, 48:556–66.

Pearlson, G., and Marsh, L. (1999). Structural brain imaging in schizophrenia: a selective review. *Biological Psych*, 46:627.

Plaut, D. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, 12(5/6).

Plum, F. (1972). Prospects for research on schizophrenia 3: Neurophysiology. *Neurosci Res Program Bull*, 10:384–8.

Pollack, J. (1990). Recursive distributed representations. *Artificial Intelligence*, 46(1):159–216.

Postle, B., Druzgal, T., and D'Esposito, M. (2003). Seeking the neural substrates of visual working memory storage. *Cortex*, 39:927946.

Potkin, S. G., Turner, J. A., Brown, G. G., McCarthy, G., Greve, D. N., Glover, G. H., Manoach, D. S., Belger, A., Diaz, M., Wible, C. G., Ford, J. M., Mathalon, D. H., Gollub, R., Lauriello, J., O'Leary, D., van Erp, T. G. M., Toga, A. W., Preda, A., Lim, K. O., and FBIRN (2009). Working memory and DLPFC inefficiency in schizophrenia: the FBIRN study. *Schizophr Bull*, 35:19–31.

Reggia, J., Ruppin, E., and Glanzman, D., editors (1999). *Progress in Brain Research*. Amsterdam: Elsevier.

Riedl, M. (2010). Case-based story planning: Creativity through exploration, retrieval, and analogical transformation. *Minds and Machines*, 20(4).

Riedl, M., Thue, D., and Bulitko, V. (2011). Game AI as storytelling. In Calero, P., editor, *Artificial Intelligence for Computer Games*. Springer Verlag.

Riedl, M., and Young, R. (2006). Story planning as exploratory creativity: Techniques for expanding the narrative search space. *Journal of New Generation Computing*, 24(3):303–323.

Ron, M., and Harvey, I. (1990). The brain in schizophrenia. *J Neurol Neurosurg Psychiatry*, 53:725–6. Comments ibid. 1992; 55: 522 and 981.

Rowe, J., Shores, L., Mott, B., and Lester, J. (2010). Individual differences in gameplay and learning: A narrative-centered learning perspective. In *Proceedings of the Fifth International Conference on Foundations of Digital Games*, 171–178. Monterey, California.

Rubin, D., and Greenberg, D. (2003). The role of narrative in recollection: A view from cognitive psychology. In Fireman, G., McVay, T., and Flanagan, O., editors, *Narrative and Consciousness*, chapter Chapter 4, 53–83. Oxford University Press.

Ruppin, E., Reggia, J., and Horn, D. (1996). Pathogenesis of schizophrenic delusions and hallucinations: A neural model. *Schizophrenia Bulletin*, 22:105–23.

Sakkalis, V., Oikonomou, T., Pachou, E., Tollis, I., Micheloyannis, S., , and Zervakis, M. (2006). Time-significant wavelet coherence for the evaluation of schizophrenic brain activity using a graph theory approach. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 1, 4265–8.

Salamone, J., Correa, M., Farrar, A., and Mingote, S. (2007). Effort-related functions of ncleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, 191(3):461–82.

Sarle, W. (1997). Neural network FAQ, part 2 of 7: Learning. Periodic posting to the usenet newsgroup comp.ai.neural-nets. URL: ftp://ftp.sas.com/pub/neural/FAQ.html.

Saugstad, L. (1994). The maturational theory of brain development and cerebral excitability in the multi-factorially inhereted manic-depressive psychosis and schizophrenia. *International Journal of Psychophysiology*, 18:189–203.

Schank, R. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 552–631.

Schank, R. (1982). *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*. Cambridge University Press.

Schank, R. (1991). *Tell Me a Story: A New Look at Real and Artificial Intelligence*. New York: Simon & Schuster.

Schank, R. (1999). *Dynamic Memory Revisited*. Cambridge, UK: Cambridge University Press, second edition.

Schank, R., and Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Erlbaum.

Schank, R., and Cleary, C. (1995). *Engines for education*. Hillsdale, NJ: Erlbaum Assoc.

Schlör, K., Moises, H., Haas, S., and Rieger, H. (1985). Schizophrenia, psychoticism, neuroleptics, and auditory evoked potentials. *Pharmacopsychiatry*, 18(5):293–6.

Schobel, S., Lewandowski, N., Corcoran, C., Moore, H., Brown, T., Malaspina, D., and Small, S. (2009). Differential targeting of the CA1 subfield of the hippocampal formation by schizophrenia and related psychotic disorders. *Archives of General Psychiatry*, 66(9):938–46.

Schott, B., Seidenbecher, C., Fenker, D., Lauer, C., Bunzeck, N., Bernstein, H., Tischmeyer, W., Gundelfinger, E., Heinze, H., and Duzel, E. (2006). The dopaminergic midbrain participates in human episodic memory formation: evidence from genetic imaging. *Journal of Neuroscience*, 26(5):1407–17.

Selemon, L., Rajkowski, G., and Goldman-Rakic, P. (1995). Abnormally high neuronal density in schizophrenic cortex: A morphometric analysis of prefrontal area 9 and occipital cortex area 17. *Archives of General Psychiatry*, 52:805–18.

Selemon, L., Rajkowski, G., and Goldman-Rakic, P. (1998). Abnormally high neuronal density in prefrontal area 46 in brains from schizophrenic patients: Applications of 3-dimensional stereologic counting method. *Journal of Comparative Neurology*, 392:402–12.

Servan-Schreiber, D., Printz, H., and Cohen, J. (1990). A network model of cathecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science*, 249:892–5.

Sibley, D., Kello, D., CT andPlaut, and Elman, J. (2008). Large-scale modeling of wordform learning and representation. *Cognitive Science*, 32:741754.

Silberman, Y., Bentin, S., and Miikkulainen, R. (2007). Semantic boost on episodic associations: An empirically based computational model. *Cognitive Science*, 31(4):645–671.

Smith, E., and Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33:5–42.

Spitzer, M. (1997). A cognitive neuroscience view of schizophrenic thought disorder. *Schizophrenia Bulletin*, 23(1):29–50.

Spitzer, M., Braun, U., Hermle, L., and Maier, S. (1993a). Associative semantic network dysfunction in thought-disordered schizophrenic patients: Direct evidence from indirect semantic priming. *Biological Psychiatry*, 34(12):864–77.

Spitzer, M., Braun, U., Maier, S., Hermle, L., and Maher, B. (1993b). Indirect semantic priming in schizophrenic patients. *Schizophrenia Research*, 11(1):71 – 80.

Spitzer, M., Kwong, K., Kennedy, W., Rosen, B., and Belliveau, J. (1995). Category-specific brain activation in fMRI during picture naming. *Neuroreport*, 6:2109–12.

Stanley, K. O., Bryant, B. D., Karpov, I., and Miikkulainen, R. (2006). Real-time evolution of neural networks in the nero video game. In *Proceedings of the Twenty-First National Conference on Artificial Intelligence*.

Stevens, J. (1992). Abnormal reinnervation as a basis for schizophrenia: A hypothesis. *Archives of General Psychiatry*, 49:238–43.

Takashima, A., Nieuwenhuis, I., Jensen, O., Talamini, L., Rijpkema, M., and Fernandez, G. (2009). Shift from hippocampal to neocortical centered retrieval network with consolidation. *Journal of Neuroscience*, 29(32):10087–93.

Tallent, K., Weinberger, D., and Goldberg, T. (2001). Associating semantic space abnormalities with formal thought disorder in schizophrenia: use of triadic comparisons. *Journal of Clinical & Experimental Neuropsychology*, 23:285–96.

Tan, H., Callicott, J., and Weinberger, D. (2007). Dysfunctional and compensatory prefrontal cortical systems, genes and the pathogenesis of schizophrenia. *Cereb Cortex*, 17 Suppl 1:i171–81.

Trantham-Davidson, H., Neely, L., Lavin, A., and Seamans, J. (2004). Mechanisms underlying differential D1 versus D2 dopamine receptor regulation of inhibition in prefrontal cortex. *Journal of Neuroscience*, 24(47):10652–59.

Tremont, G., Halpert, S., Javorsky, D., and Stern, R. (2000). Differential impact of executive dysfunction on verbal list learning and story recall. *Clinical Neuropsychologist*, 14:295–302.

van Os, J., and Kapur, S. (2009). Schizophrenia. *Lancet*, 374:635–45.

van Rossum, G., and de Boer, J. (1991). interactively testing remote servers using the python programming language. *CWI Quarterly*, 4(4):283–303.

Vinogradov, S., King, R. J., and Huberman, B. A. (1992). An associationist model of the paranoid process: Application of phase transitions in spreading activation networks. *Psychiatry*, 55:79–94.

Wang, H., and Fan, J. (2007). Human attentional networks: A connectionist model. *Journal of Cognitive Neuroscience*, 19:1678–89.

Wechsler, D. (1987). *Manual for the Wechsler memory scale revised*. The Psychological Corporation, San Antonio, TX.

Weinberger, D. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. *Archives of General Psychiatry*, 44(7):660–9.

Weinberger, D. (1995). From neuropathology to neurodevelopment. *Lancet*, 346:552–7.

Wendling, F. (2008). Computational models of epileptic activity: a bridge between observation and pathophysiological interpretation. *Expert Rev Neurother*, 8(6):889896.

Winder, R., Cortes, R., Reggia, J., and Tagamets, M. (2007). Functional connectivity in fMRI: A modeling approach for estimation and for relating to local circuits. *Neuroimage*, 34:1093–107.

Winterer, G., and Weinberger, D. (2004). Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Tr Neurosci*, 27:683–90.

Wittmann, B., Schott, B., Guderian, S., Frey, J., Heinze, H., and Duzel, E. (2005). Reward-related fMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*, 45(3):459–67.

Woods, S. (2003). Chlorpromazine equivalent dosages for the newer atypical antipsychotics. *J Clin Psychiatry*, 64:663–667.

Young, R. (2007). Story and discourse: A bipartite model of narrative generation in virtual worlds. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 8(2):177–208.

Young, R., Riedl, M., Branly, M., Jhala, A., Martin, R., and Saretto, C. (2004). An architecture for integrating plan-based behavior generation with interactive game environments. *The Journal of Game Development*, 1(1):51–70.